



Reducing Affective Responses to Surgical Images and Videos Through Stylization

Lonni Besançon, Amir Semmo, David J. Biau, Bruno Frachet, Virginie Pineau, El Hadi Sariali, Marc Soubeyrand, Rabah Taouachi, Tobias Isenberg, Pierre Dragicevic

► To cite this version:

Lonni Besançon, Amir Semmo, David J. Biau, Bruno Frachet, Virginie Pineau, et al.. Reducing Affective Responses to Surgical Images and Videos Through Stylization. Computer Graphics Forum, 2020, 39 (1), pp.462–483. 10.1111/cgf.13886 . hal-02381513v2

HAL Id: hal-02381513

<https://inria.hal.science/hal-02381513v2>

Submitted on 12 Dec 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Reducing Affective Responses to Surgical Images and Videos through Stylization

Lonni Besançon,¹ Amir Semmo,² David Biau,³ Bruno Frachet,³ Virginie Pineau,⁴ El Hadi Sariali,³ Marc Soubeyrand,³ Rabah Taouachi,⁴ Tobias Isenberg,⁵ and Pierre Dragicevic⁵

¹Linköpings Universitet, Norrköping, Sweden

²Hasso Plattner Institute, University of Potsdam, Germany

³Assistance Publique – Hôpitaux de Paris, France

⁴Institut Curie, France

⁵Inria, France



Original photo of a lasagna dish.

Output of HUESHIFT2.

Output of FLOWABS [KD08].

Figure 1: Two techniques studied in this article, each using a different strategy for making surgery images easier to look at.

Abstract

We present the first empirical study on using color manipulation and stylization to make surgery images/videos more palatable. While aversion to such material is natural, it limits many people's ability to satisfy their curiosity, educate themselves, and make informed decisions. We selected a diverse set of image processing techniques to test them both on surgeons and lay people. While color manipulation techniques and many artistic methods were found unusable by surgeons, edge-preserving image smoothing yielded good results both for preserving information (as judged by surgeons) and reducing repulsiveness (as judged by lay people). We then conducted a second set of interview with surgeons to assess whether these methods could also be used on videos and derive good default parameters for information preservation. We provide extensive supplemental material at osf.io/4pfes/.

CCS Concepts

• **Computing methodologies** → **NPR**; Image processing; • **Human-centered computing** → Empirical studies in HCI;

1. Introduction

Many non-photorealistic and expressive rendering techniques deal with the stylization of 2D images or videos [KCW13, RC13]. While much of this work was initially motivated by the desire to replicate artistic techniques and was only guided by a subjective visual comparison to existing artwork, researchers have begun to empirically evaluate the effects of stylization [Ise13, Sal02, GLJ*10]. Some researchers argue, however, that controlled experiments are difficult in the context of expressive rendering [Mou14], and that we should rather concentrate on subjective evaluation [Mou14] and

on the appreciation of resulting graphics [HL13]. While such forms of evaluation arguably have their place in the context of the often art-inspired field of expressive rendering, the goal of creating expressive graphics is increasingly understood to incorporate more than the “support of artists (or illustrators)” or the “creation of tools for visual expression for non-artists” and to include, e. g., also “illustrations [...] to inform [...] patients” in a medical context [Ise16]. In this latter case it is then essential that we understand how different stylistic filters are perceived and experienced by real people, and that we thus study them empirically, through controlled experiments [GLJ*10].

In the past, in fact, some researchers have already examined such effects of stylization. Mandryk et al. [MML11] resp. Mould et al. [MML12], e. g., studied people's emotional response to stylized images and found that emotional responses were generally muted, and that responses concentrated around neutral feelings. Others [DWI13] argue (motivated by the results of earlier studies [SSRL96]) that stylization may affect people's attitude toward a data visualization and result in longer times they spend looking at the visuals. Here, however, we are less interested in potentially positive effects of stylization, but instead in how much it can diminish negative affects caused by unpleasant pictures. Such pictures are involved when surgeons inform their patients before surgical procedures—because many people find surgery pictures repellent (e.g., [TSL97, SLW*02]), effective communication can suffer. This context would seem like an ideal application case for expressive rendering [Ise16]. Nonetheless, the creation of effective illustrative visualizations of a wide variety of surgical procedures is still beyond our abilities. We thus study whether it is possible to use existing stylization techniques for 2D images—applied to real surgical images—to achieve a similar effect, and diminish the negative affect that surgery pictures can elicit. Applications go beyond patient information and include student training (some medical students will not become surgeons), media communication, and public education.

This article is an extended version of our previous paper [BSB*18], published at Expressive 2018. As in the original paper, we first discuss various image filtering and stylization techniques that can be used to dampen the negative affect elicited by surgery pictures. We then report on an interview session with four surgeons, who helped us differentiate between techniques that can preserve important information, and techniques that are unusable because they obfuscate too much. We then report on an experiment where the most promising techniques were tested on ordinary subjects. We found that all techniques can reduce the repulsiveness of surgery pictures as judged by participants, although spatial-domain techniques appear to be more potent than color manipulations. We then extend this work with an experiment to find suitable parameters for a given stylization technique to support the intended degree and style of abstraction, a discussion of how to extend the technique to video, results from a second set of interviews with surgeons based on the improved implementation, as well as an implementation of the technique for use in the browser to filter offensive content. We conclude by a discussion and opportunities for future work.

2. Background

In this section, we review related work in non-photorealistic rendering, before reviewing work from other areas on how people perceive, experience, and deal with surgery and injury images.

2.1. Non-photorealistic Surgery Illustrations

Medical illustration has long been among the primary motivations for non-photorealistic and expressive rendering [GG01, SS02] and, consequently, many researchers have developed rendering techniques for this purpose. Several surveys and tutorials cover the field in detail (e. g., [CSESS05, ECS06, VCSE*06, PB14, LVPI18]) and we thus refrain from citing specific techniques. Common among

them is, however, that they are inspired by traditional, usually hand-made illustration techniques, styles, and examples—they thus focus on clarification and explanation, rather than on emotion or on reducing the negative affect that certain content could induce in people.

Another common characteristic of many illustrative techniques and also traditional illustration styles—for medical application and otherwise—is the use of abstraction and emphasis. These aspects have been discussed in the visualization and expressive rendering literature, such as in the contributions by Rautek et al. [RBGV08] and Viola and Isenberg [VI18]. Here, abstraction is “*a transformation which preserves one or more key concepts and removes detail that can be attributed to natural variation, noise, or other aspects that one intentionally wants to disregard from consideration*” [VI18]—to allow viewers of a visualization to focus on major or important aspects. In this work, however, we explore the abstracting qualities of image filters for the removal of details such that the images are perceived as less offensive—potentially because they no longer depict surgery situations in all their details.

2.2. Non-photorealistic Techniques and Affect

In the past, researchers have studied how stylization can influence how people perceive images. Gooch and Willemsen [GW02], e. g., showed that a line-based rendering of a virtual scene leads participants to underestimate distances by about a third, quite similar to what happens in ‘photorealistic’ virtual reality (VR) settings. Later, Gooch et al. [GRG04] showed that non-photorealistic illustrations and caricatures of people's portraits could be learned faster than real photographs. We cannot deduct from these results, however, that stylized images would lead people to feel differently about what is shown.

Already early work on non-photorealistic rendering, however, discussed this very effect. Duke et al. [DBHM03] and Halper et al. [HMH*03], for example, described how the (non-photorealistic) depiction style can affect people's assessment of danger and safety as well as strength and weakness, and can change their participation and interaction behavior (for study details see Section 2 of Halper's thesis [Hal03]). Even before this work, Schumann et al. [SSRL96] provided evidence for stylization to increase people's willingness to interact with visuals. More recently, McDonnell et al. [MBB12] showed that an increased abstraction of virtual characters (according to their participants' classification of “realism”) decreases appeal, friendliness, and trustworthiness up to a point; for highly abstracted depictions people again feel similar about the stylized virtual characters as they do for realistic depictions—similar to what the Uncanny Valley theory predicts. Like the perceptual studies discussed before, however, these approaches do not shed light on whether stylization changes people's negative emotions toward disturbing images.

Most relevant for our own work, out of the expressive rendering literature, is Mandryk et al.'s [MML11] and Mould et al.'s [MML12] work who demonstrated that stylization can affect the emotional interpretation of images. Similar to what we do in our experiment, they applied a range of styles (stippling, line art, painterly rendering, and blur) to a set of images with different affective content from the International Affective Picture System (IAPS), and analyzed people's feeling of arousal, valence, dominance, and aesthetics.

Stylization generally muted participants' emotional responses toward a neutral point, yet emotions were never completely suppressed. Their negative stimuli (e. g., a gun pointed at camera, or a cemetery), however, did not have the repulsive potency that surgery photos can have. This study thus inspires our own, but we specifically target surgery pictures that many people cannot easily look at.

We note that researchers also have examined the opposite path: changing the stylization of images based on emotions detected in a video feed. Shugrina et al. [SBC06], for example, presented their “empathic painting” technique that recognizes a person's emotional state based on features of their facial expression, which they then use to adjust the parameters of a painterly rendering technique. Here, Shugrina et al. borrow from the psychological literature and created a mapping from the detected emotional state to rendering parameters such as stroke path and color. Yet, it is not clear if the resulting images also change the emotional state of the viewer, or if so in what way this process can be controlled.

2.3. Human Response to Surgery Imagery

For several decades, researchers have studied human response to repellent images to uncover the physiological and psychological mechanisms involved. Studies have used various types of aversive stimuli such as homicide scenes [HWBS70], spiders [TSL97], vomit [OHMD08], maggots, cadavers, and dirty toilets [SISW*02]. Closer to our concerns, many studies have examined responses to scenes depicting a body envelope violated by an injury or a surgery. Examples include photos of body mutilation (e. g., [KWW77]), of surgery procedures (e. g., [TSL97, SLW*02]), and videos of medical interventions such as blood draw [GD12], open-heart surgery [OHMD08], or surgical amputation [RH08]. Studies have involved both ordinary subjects (e. g., [HWBS70]), BII-phobic[†] subjects (e. g., [ÖSL84]), and often a combination of both (e. g., [HS14]).

Researchers have employed various measurements to quantify subject reactions, the most common being heart rate (e. g., [KWW77, OHMD08]). Others include facial expression using videotaping [LM92] or electromyography [LGBH93, OHMD08], skin conductance [LGBH93], neural activation using fMRI [SISW*02], attentional avoidance using eye tracking [AHO13], and visuomotor processing using a response priming task [HS14]. Researchers also used subjective measures, asking subjects to report to what extent they felt fear and disgust [TSL97, SLW*02], avoided watching [OHMD08], or experienced vasovagal (i. e., pre-fainting) symptoms [GD12]. A strong reaction to a body injury depiction is often marked by a decrease in heart rate, or an increase followed by a rapid decrease called “diphasic response” [COL09]. It also often involves activation of the corrugator supercilii (the “frowning muscle”) and the levator labii (which lifts the upper lip) [COL09]. However, studies are inconsistent and there appears to be no perfectly reliable measure that can consistently elicit the same response [COL09].

[†] Blood Injection and Injury phobia (BII phobia) refers to “an extreme and irrational fear of blood, injuries, or of receiving an injection or an invasive medical procedure” which affects about 3.5% of the population [HS14].

Most of these studies were conducted to untangle the emotions involved when people witness surgeries or injuries, sometimes in the hope of better treating BII phobia. This has proven hard to study, as reactions seem to involve various emotions such as anxiety, fear, disgust, and vicarious pain [COL09, BLD*08]. In particular, the relative role of fear vs. disgust has long been a subject of debate, although now the consensus seems to be that disgust is the main emotion involved [COL09, OCMP10]. To understand why, it helps to recall that fear has evolved for organisms to run away from threats such as spiders, but for static content like body injuries, no such response is necessary [COL09]. More likely, body injuries are experienced as repellent to prevent risks of disease or contagion following physical contact, which requires a disgust response [COL09]. Chapman and Anderson [CA12] introduced a taxonomy of disgust where *blood-injury disgust* is a subtype of *physical disgust*, and whose evolutionary function is to avoid infection. Olatunji et al. [OHMD08], however, distinguishes *contamination disgust* from *animal-reminder disgust*, with animal-reminder disgust being elicited by “attitudes and practices surrounding sex, injury to the body or violations of its outer envelope, and death” which all act as “reminders of our own mortality and inherent animalistic nature” [OHMD08].

Despite all this previous work, human reactions to the sight of surgery scenes remains poorly understood. Our goal is not to further this understanding, but simply to find out whether processing surgery photos can dampen their affective potency. As far as we know, all studies on blood-injury disgust have either assessed aversive stimuli in isolation or compared them with neutral stimuli, and none of them has studied the effect of processing aversive stimuli using filters or stylization. When conducting our study, we drew from the experience accumulated in this research area, but simplified the methods to directly answer our research question.

In parallel to this body of work focusing on blood-injury disgust, there has been work in psychology and the neurosciences where various types of emotionally-salient stimuli were used to study emotion and cognition. Such stimuli were used, for example, to study cultural differences in emotion processing [WKG*03], and emotion regulation [EVW*07]. Some of the stimuli involved surgery and injury photos but again, affective neutralization through image processing has not been a focus. Nevertheless, this area of research has produced standardized stimuli sets which we will use for our own study, as explained in Section 5.1.

2.4. Picture Censorship Practices

On a societal level, offensive imagery has been addressed in two major ways: legal censorship and *de facto* (or self) censorship. While there appears to be close to no legal restriction on what visual content can be published in newspapers [Too14] or in Wikipedia [Wik10a, Wik10b], films and video games are usually regulated by rating systems to classify the media with regard to its suitability for different audiences. While movies cannot be easily customized, the video game industry has explored a wide range of “adjustable censorship” techniques. Some old video games had violent and sexual content disabled by default, while giving the option to reactivate it through the use of secret codes. More elaborate adjustable censorship techniques were also developed: some video games (e. g.,

Silent Hill, Resident Evil, House Of The Dead, later release of Ocarina of Time) offer the option to change the color of blood to various tones such as blue, dark, or green depending on the game [TVT18a, Ze18]. While most mangas feature black blood due to the constraints of black & white printing, some color animes and animated films employ a different blood color to suit all audiences (e. g., Dragon Ball Kai, Pokémon, Bleach, The little mermaid) [TVT18b]. Similarly, in movies, black and white has been occasionally used to censor scenes with excessive bloodshed [Kil17]. All such practices suggest that blood is considered to epitomize violence, but once deprived from its characteristic red color it seems to suddenly become inoffensive in people's minds. These practices provide motivation for considering simple color manipulation techniques in our work.

3. Choice of Processing Techniques

In this article we use the term *image processing technique* or simply *technique* to refer to any procedure that transforms an image into another image, while keeping it recognizable. We considered four classes of techniques of varying complexity: color manipulation, edge-preserving smoothing, edge detection and enhancement, and image-based artistic rendering. We first outline relevant work and provide rationales for the techniques we retained, and then provide parameter settings yielding reasonable levels of abstraction. For the selection of the techniques, we strove for classical or state-of-the-art methods that cover the taxonomy proposed by Kyprianidis et al. [KCWI13]. Further, the techniques need to process surgery pictures in a content-preserving way, because we want to target applications where negative affects are diminished but the abstract pictures can still be used for patient information or media communication. The selected 13 techniques and their settings are summarized in Table 1 and illustrated in Figures 1 and 2 with a 1024×768 photo.

We intentionally did not include techniques that require user input such as locations of particular focus because the need for manual input would limit the possible application domains of our work. We thus restricted our exploration to methods applied to the whole image equally. This means we excluded, for example, focus-and-context methods such as lens-based distortion (e. g., [CCF97]) or interactive stylization (e. g., [Hae90, CAS*97, SIMC07]). We also excluded techniques that would drastically reduce the amount of detail included in the images such as global pixelization (e. g., [GDA*12, IK12])—local pixelization again would require user input. Such approaches could be examined as future work.

3.1. Color Manipulation

We considered two approaches for changing blood color: decolorization and recolorization [Pra07].

Decolorization. Grayscale conversion is a popular method for image decolorization, where the main challenge is to preserve and make use of the chrominance components so that perceptual image features are retained [Čad08, MZZW15]. Most algorithms transform the problem into optimization to preserve salient features, e. g., by quantifying color differences between image locations [GOTG05] or prevailing chromatic contrasts [GD07], optimizing color and luminance contrasts [NvN07], or considering the

Helmholtz-Kohlrausch color appearance effect [SLTM08]. The latter localized apparent grayscale algorithm performed best in a previous experiment [Čad08] and we thus retained it and named it APPARENTGREY in this article. However, the method may suffer non-homogeneity artifacts near region boundaries, which can be addressed with a global mapping scheme [KJDL09].

Recolorization. A simple yet effective method of recolorization is to hue shift the colors in hue-saturation-value (HSV) space. We consider a uniform hue shift which makes blood appear in a different color, to which we refer to as HUESHIFT. The hue shift shown in Figure 1 uses different settings and we discuss it in Section 5.2. A more sophisticated approach could involve color transfer between source and target images or color palettes, which typically relies on image statistics to globally and locally control color distributions [FPC*14].

3.2. Edge-preserving Image Smoothing

While color manipulation may reduce the emotional impact carried by blood, it preserves the details of the original photo. A black-and-white photo, in particular, may still appear too crude. We thus consider other types of filters, starting with edge-aware image smoothing as a building block for abstraction, artistic stylization, and tone mapping. Numerous automatic filter-based techniques have been proposed for these applications, typically by approximating an anisotropic diffusion [Wei99], i. e., to smooth details without filtering significant image structures. The balancing between both aspects varies between filters and is thus critical for us. Consequently, we favored filter enhancements that derive local image structures for improved feature-aware processing by adapting the filter kernels to the shape, scale, and orientation of the local image structure.

Bilateral filter. The bilateral filter is a popular choice to approximate an anisotropic diffusion by weight-averaging pixel colors in a local neighborhood based on their distances in space and range [TM98]. It weights pixels with a high difference in intensity less than a Gaussian filter to preserve image structures at a better scale. We retain it and refer to it as BILATERAL. Most relevant applications apply the bilateral filter in a multi-stage process for real-time rendering with a cartoon look [WOG06], and enhance it by flow-based implementations adapted to the local image structure [KD08, KLC09], in particular to reduce smoothing across falsely detected edges. Providing smooth outputs at curved boundaries of delicate structures, we thus consider the flow-based variant [KD08], and name it FLOWABS. As a generalized variant, the guided filter [HST13] may provide similar characteristics with reduced unwanted gradient reversal artifacts, but only provides a non-feature-aligned implementation.

Mean-shift. A mean-shift is a popular approach for edge-preserving smoothing [CMM02] and saliency-guided image abstraction [DS02]. It provides a non-parametric filter that estimates probability density functions by iteratively shifting color values to averaged color values of a local neighborhood. However, the approach typically produces rough boundaries that is more suited to image segmentation.

Kuwahara filter. A popular approach that works accurately even

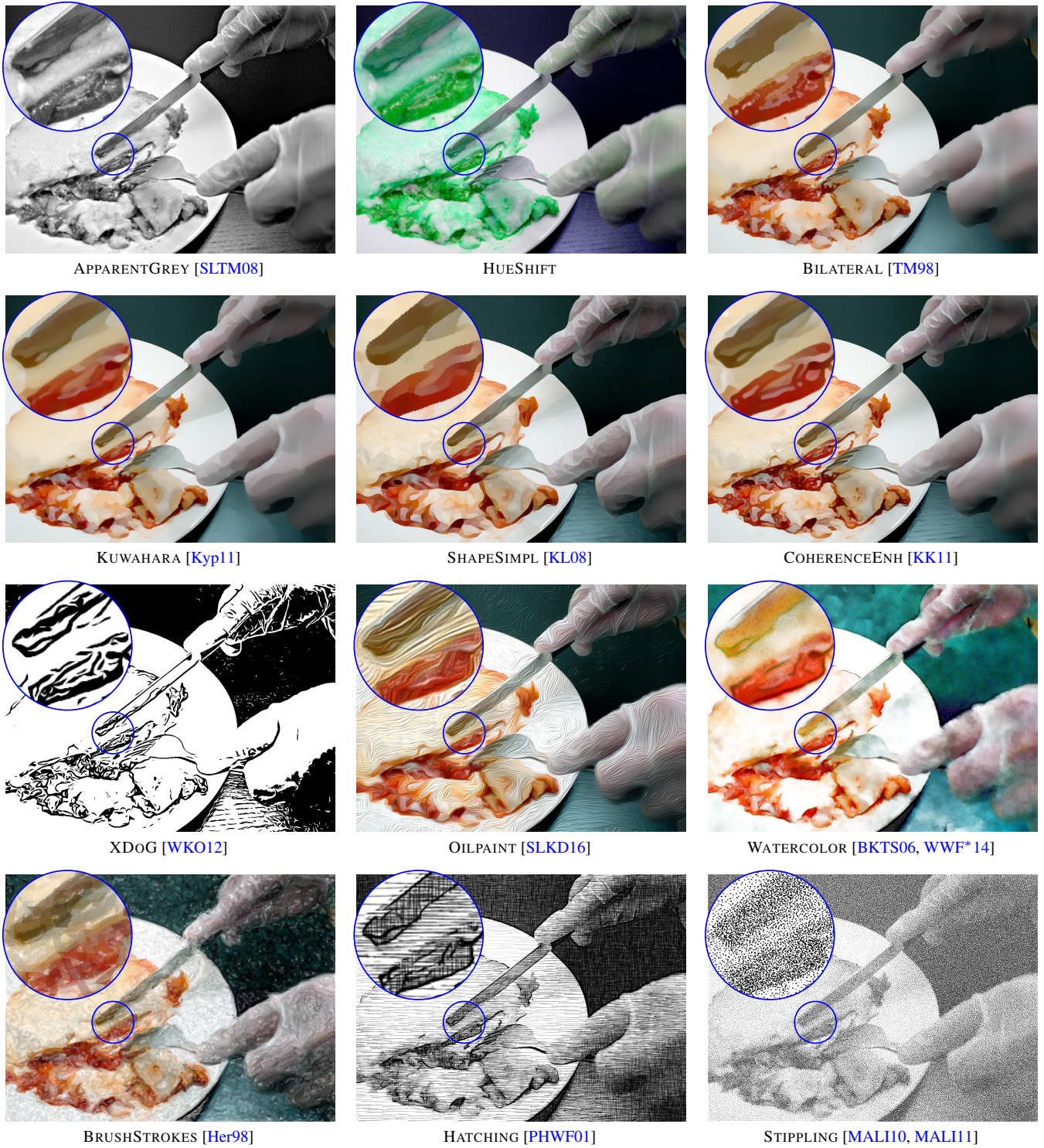


Figure 2: Image processing techniques used in the interviews with surgeons, together with FLOWABS shown in Figure 1.

with high-contrast images—contrary to the bilateral filter—and provides smoothed outputs at curved boundaries, is the Kuwahara filter [KHEK76] and its generalized [PPC07] and anisotropic [KKD09, Kyp11] variants. The kernel of the anisotropic Kuwahara

filter is divided into shape-aligned overlapping subregions, where the response is defined as the mean of the subregion with minimal variance. We retain the multi-scale variant [Kyp11] and refer to it

Table 1: Overview of techniques with their correspondent parameters.

	Abbreviation	Based on works	Parameter settings
Filters	APPARENTGREY	[SLTM08]	$N = 4, p_i = 0.5, k_i = 0.5$
	HUESHIFT	custom (Section 3.5)	$\alpha = 120.0$
	BILATERAL	[TM98]	$\sigma_d = 3.0, \sigma_r = 4.25\%$
	FLOWABS	[KD08]	$\rho = 2, n_e = 1, n_d = 3, \sigma_d = 6, \sigma_r = 5.25\%, \sigma_e = 1, \tau = 0.99, \epsilon = 0, \phi_e = 2, \sigma_m = 3, q = 8, \phi_q = 2$
	KUWAHARA	[Kyp11]	$N = 8, \rho = 2.0, r = 6, q = 8, \alpha = 1, \tau_w = 0.02, p_s = 0.5, p_d = 1.25, \tau_v = 0.1$
	SHAPESIMPL	[KL08]	$N = 8, k = 8, r = 1, \text{etf}_{\text{halfw}} = 3, \text{etf}_N = 4, \sigma_{\text{shock}} = 1, \tau_{\text{shock}} = 1$
	COHERENCEENH	[KK11]	$N = 4, \sigma_d = 1, \tau_r = 0.002, \sigma_i = 6, \sigma_g = 1.5, r = 2, \tau_s = 0.005, \sigma_a = 1.5$
	XDOG	[WKO12]	$\sigma_c = 2.28, \sigma_e = 1, \sigma_m = 4.4, p = 99.0, \phi = 100.0, \epsilon = 0.65, \sigma_a = 1.0$
Stylization	OILPAINT	[SLKD16]	$\sigma_s = 12.0, n_e = 0, \sigma_b = 12.0, k_{\text{scale}} = 5.0, k_{\text{specular}} = 0.8, k_{\text{shininess}} = 12.0$
	WATERCOLOR	[BKTS06, WWF*14]	Effects of [BKTS06, WWF*14] using FLOWABS: $\rho = 2.0, \sigma_d = 4.0, \sigma_r = 15.00\%$
	BRUSHSTROKES	[Her98]	$T = 200, \text{size}_{\text{max}} = 8, f_c = 1, f_\sigma = 0.5, j_{\{h,s,v\}} = 0, j_{\{r,g,b\}} = 0.3, f_g = 1, \sim = 0.5, \text{length} = [4, 16]$
	HATCHING	[PHWF01]	Art map of [PHWF01] scaled at 0.6, linear mapping with luminance, edge settings of FLOWABS
	STIPPLING	[MALI10, MALI11]	$\text{res}_0 = 1200 \text{ ppi}, f_p = 2.0, \text{placement randomness} = 25\%, \text{distribution} = \text{normal}, \text{colors} = b\&w, \tau = 127$

as KUWAHARA. It maintains a uniform level of abstraction due to local area flattening and can scale with the image resolution.

Shock, morphological, and geodesic filters. Contrary to previous filters, additional categories weight colors across feature boundaries for higher levels of abstraction, for which we retain methods with shock filtering, i. e., in conjunction with a constrained mean curvature flow [KL08] (SHAPESIMPL) and diffusion tensors for coherence-enhancing abstraction [KK11] (COHERENCEENH). Morphological filtering based on dilation and erosion, and geodesic filtering using distance transforms are also popular choices to obtain results of high abstraction [CSR10, Mou12], but were found to require local control to effectively adjust the level of abstraction.

Filters using global optimizations. Many filters focus on image decompositions by solving optimization problems to separate detail from base information, e. g., based on weighted least squares [FFLS08], histograms [KS10], and gradient minimization [XLXJ11]. While they have strengths in applications requiring complementary global optimizations such as tone mapping and colorization, they are typically not suited for interactive applications.

3.3. Edge Detection and Enhancement

Winnemöller et al. [WKO12] distinguish between gradient-based edge detection that thresholds the gradient magnitude of an image and Laplacian-based edge detection that identifies zero-crossings in the second derivative. Popular gradient-domain approaches identify image gradients with high magnitudes by using convolution filters, such as the Prewitt and Sobel filter [Pra07], with subsequent thresholding of the magnitude. The approach is popular with medical images to ease object recognition, however produces results that are sensitive to noise. The Canny edge detector [Can86] as a multi-stage algorithm provides several enhancements by combining smoothing and differentiation operators. However, although popular with magnetic resonance imaging (MRI) and computed tomography (CT) images, it is more directed to semantic segmentation and may produce disconnected edge segments.

A real-time approach that is less sensitive to noise is to approximate the Laplacian of Gaussian [MH80] using difference of Gaussians (DoG). The approach has shown to provide smooth edges of delicate structures, e. g., with respect to human faces [GRG04],

since they are adapted to the local orientation of an input image to create smooth coherent outputs for line and curve segments. Counterbalancing strong simplifications of the bilateral filter, we thus combine it with the enhanced separable flow-based implementations of the DoG [KLC07, KD08, WKO12] for FLOWABS, to reintroduce filtered structures as enhanced visual cues and obtain a cartoon-like effect. We also retain the XDOG filter [WKO12] as a generalized approach that is able to obtain two tone black-and-white images, which relates to drawings found in illustrative visualization.

3.4. Image-based Artistic Rendering

Artistic image stylization has been suggested to dampen emotional responses [MML11]. We considered stylization techniques that simulate traditional media and painting techniques found in illustrative visualization, i. e., watercolor, oil paint, pen-and-ink, and stippling. Image filters are prominently used as building blocks of complex stylization effects, such as the bilateral filter and DoG to obtain toon renderings (FLOWABS), and flow-based Gaussian smoothing for more abstract filtering that simulates oil paint [SLKD16] (OILPAINT). In addition, we use a WATERCOLOR technique that simulates effects such as pigment density variation, edge darkening, wet-in-wet, and wobbling [BKTS06, WWF*14]. For stroke-based rendering, a popular method is to iteratively align brush strokes of varying color, size, and orientation according to the input image, for which we use Hertzmann's [Her98] approach (BRUSHSTROKES). Techniques for tonal depiction typically direct tonal art maps based on luminance, for which we use a 2D hatching implementation that borrows from Praun et al. [PHWF01] coupled with a DoG-based edge (HATCHING). Finally, we consider the example-based stippling technique described by Martín et al. [MALI10, MALI11] (STIPPLING), as it is able to provide scale-dependent results.

3.5. Parameters and Configurations

The parameterization of the selected techniques is a crucial aspect to strike a balance between filtering information and retaining significant image structures. We thus based our default configurations on the presets reported by the original authors of each technique, under the assumption that their findings on generalized photos can also be applied to surgical images if details and properties such as contrast

are balanced. In addition, we followed the approach by Tomasi and Manduchi [TM98] of applying multiple iterations of the bilateral filter—and Gaussian-based filters in general—to preserve edges at a better scale while still being able to provide strong simplification effects. The parameters and configurations summarized in Table 1 are based on the reported results and default presets, and used in our experiments for medical images at a resolution of 1024×768 pixels. For images whose resolution differ from 1024×768 , the techniques' parameters that relate to spatial distances can be linearly scaled to obtain a stable level of abstraction. In the following, we briefly explain the rationale behind these configurations and kindly refer to the respective original works for in-depth discussions.

Filters. We set the APPARENTGREY filter to use default settings with uniform spatial control of four subbands to locally adjust local chromatic contrasts. For HUESHIFT, we use a shift of -120.0 degrees on the *hue* channel mapped to the hue-saturation-lightness (HSL) color wheel. We align the BILATERAL filter to obtain a soft Gaussian smoothing with a spatial distance $\sigma_d = 3.0$ using additional filtering in the CIE-Lab color space using an increased weight ($\sigma_r = 4.25\%$). For FLOWABS, we use default parameters for edge enhancement with a doubled distance for bilateral filtering to compromise with the 512×512 pixel images used by Kyprianidis and Döllner [KD08]. We use the KUWAHARA filter in a typical configuration with a radius of six pixels aligned to eight sectors, a slightly smoothed structure tensor, and multi-scale estimation [Kyp11]. For deliberate smoothing across shape boundaries using SHAPESIMPL and COHERENCEENH, we configure these filters to perform a single step of shock filtering after every iteration of mean curvature flow—i. e., four [KK11]) and eight [KL08] steps in total, respectively. Finally, the edge enhancement uses XDoG to output fine coherent lines with high details and a two-tone thresholding to sparsely obtain negative edges [WKO12].

Artistic Styles. Here we mainly seek to replicate the level of abstraction targeted by the respective works: for OILPAINT, smoothing parameters with a light paint texture that fall into the medium range as described by Semmo et al. [SLKD16]; an implementation for WATERCOLOR using flow-based bilateral filtering of FLOWABS but with wider filter kernels to achieve a similar level of abstraction as Bousseau et al. [BKTS06] and Wang et al. [WWF*14]; the “colorist wash” preset of BRUSHSTROKES defined by Hertzmann [Her98] to produce semi-transparent layered brush strokes; the default art map used by Praun et al. [PHWF01] that is linearly mapped to the luminance in CIE-Lab color space; and default parameters for STIPPLING described by Martín et al. [MALI11] at highest resolution, with a normal distribution and black-and-white thresholding.

4. Interviews with Surgeons

To determine which of these 13 techniques are useful in practice, we interviewed four surgeons: two otolaryngologists (*S1* and *S2*; 13 and 35 years experience, resp.), one orthopedic surgeon (*S3*; 10 years experience), and one reconstructive surgeon (*S4*; 10 years experience). While surgeons naturally have a high tolerance for surgery images and thus may not be able to assess whether a technique can reduce the affective response to an image, their expertise is needed to study which of the techniques can preserve information well. Although

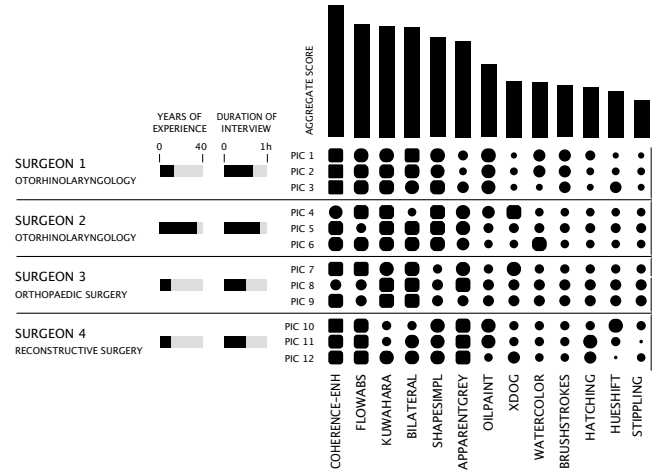


Figure 3: Visualization [PDF14] of people's preferences for each combination of surgeon, photo, and technique. Techniques are sorted according to their aggregate score, from most to least preferred.

all four surgeons are co-authors of this article, none of them was involved in the research at the time of the interviews.

We asked each surgeon to send us three of their own surgery photos that could help them explain a specific procedure to non-experts. All photos had landscape orientation and we cropped each to an aspect ratio of $4/3$, and then resized them to 1600×1200 pixels. Having all photos share the exact same pixel dimensions allowed us to better control the experiment, as the effect of a filter typically depends on the resolution of the input image. We then processed the photos as described in Section 3.5, and printed each on a separate A4 sheet. At the start of each session, we asked the surgeon to compare and classify the processed images by making piles based on how useful the image would be as a support for communication and explanation, especially in terms of how much important information is preserved. Two images perceived as equally useful would go onto the same pile. We repeated this procedure three times, once for each photo. Each surgeon saw $3 \text{ photos} \times 13 \text{ techniques} = 39$ images, in addition to the 3 original photos. To limit ordering effects, we randomized the order in which processed images were presented. After the classification, we asked the surgeons to comment further on the techniques and inquired them about possible applications.

4.1. Reported Preferences

We processed self-reported preferences as follows: first, for each combination of surgeon \times photo, we assigned a number to each technique according to its pile: 1 for the most preferred pile, 2 for the second pile, etc. We then normalized these ranks using the halfway accumulative distribution [JS04]. This method gives each rank a score between 0 and 1 that corrects for possible differences in the way ranks are assigned (e. g., when some surgeons make more piles than others). We then derived *preference scores* by inverting the normalized ranks ($y = 1 - x$). We show all preference scores in Figure 3. Finally, we averaged preference scores across pictures and surgeons to derive a single *aggregated preference score* per filter which we also show in Figure 3, on top. As we can see, COHERENCEENH was

the most preferred technique, followed by FLOWABS, KUWAHARA, and BILATERAL. We discuss the surgeons' comments next.

4.2. Qualitative Feedback

We encouraged surgeons to voice comments both during the classification and in the debriefing interview. In general, they reported that some of our processed images could be used for textbooks or classes ($\times 2$ surgeons), to communicate with patients ($\times 2$), and surprisingly even to communicate with other experts ($\times 1$), as “drawings can be augmented” with notes for instance. One surgeon reported they could be particularly useful for plastic surgery and that it would be interesting to see how automatic processing could help communicate with children patients, as they are more sensitive to surgery images. We now report on the comments that were made for more than one photo and/or by more than one surgeon.

To begin with, the BILATERAL technique was found to be usable for patients or in a book ($\times 2$ photos, 1 surgeon). In addition to its good ranking, the FLOWABS technique's resemblance to cartoons or comic strips was pointed out, and with it the possibility to remove a bit of “violence” from the photo ($\times 2$, 1). It was also praised for the high visibility it gave to contours ($\times 3$, 2). In contrast, the HATCHING technique was reported to remove too many details ($\times 5$, 3). Similarly, while HUESHIFT only manipulated colors, it was reported to cause loss of information ($\times 3$, 1), and made it hard to find anatomical correspondences ($\times 2$, 1). The OILPAINT technique generated mixed reactions. On the one hand, it was praised for its artistic look ($\times 3$, 2) and could potentially be used in books or with patients ($\times 2$, 1). On the other hand, it was reported to remove useful information ($\times 2$, 2). BRUSHSTROKES was also reported to be artistic and possibly useful with patients ($\times 2$, 2). The same qualities were reported for the SHAPESIMPL technique. STIPPLING was explicitly reported as not usable ($\times 3$, 1) and causing too much loss of information ($\times 7$, 3). The WATERCOLOR technique also removed too many details ($\times 5$, 2). Finally, XDOG was also found to cause too much information loss ($\times 2$, 1) as it makes it difficult to distinguish colors and contours ($\times 5$, 2).

5. Experiment with Lay People

Our interview with surgeons helped us understand which processing techniques preserve key information from surgery photos. However, it is hard for surgeons to accurately predict the affective impact of processed photos on lay people. Thus, we conducted an experiment where we presented surgery photos to 30 participants, both unprocessed and processed, and asked them to rate them according to how repulsive they are. This experiment was approved by Inria's ethics committee (COERLE, approval number 2017-015).

5.1. Pictures

Throughout this section, we use the term *picture* to refer to an original photo, and the term *stimulus* to refer to a processed or unprocessed picture that is meant to be presented to participants.

We selected five pictures from two research catalogs: the Nencki Affective Picture System (NAPS) [MŽJG14] and the International Affective Picture System (IAPS) [LBC97]. These catalogs contain

a range of emotionally-evocative photos that have been validated to elicit a positive, neutral, or negative affect. We selected five surgery photos among the negative pictures:[‡]

- People 202 (NAPS)—a leg surgery,
- People 216 (NAPS)—a leg surgery or autopsy,
- People 221 (NAPS)—a surgery in the eye area,
- 3212 (IAPS)—a surgery performed on a dog, and
- 3213 (IAPS)—a finger surgery.

These photos were selected as follow. For the NAPS catalog, we selected all photos in the “People” category that had a horizontal orientation, leaving us with 204 photos out of the 1356 initial ones. Ten of them were of surgical procedures, of which we selected three that involved large incisions on recognizable body parts (legs and eye). The other photos either had small incisions, or unrecognizable body parts. For the IAPS catalog, which only had horizontal photos, four photos out of the 1182 were tagged “surgery”, of which we selected two with large incisions on recognizable body parts (chest and hand). We reasoned that photos with recognizable body parts were emotionally more potent. Both catalogs came with data on average subject ratings across different emotional scales including disgust, but since data was missing for some photos in the NAPS catalog, we decided not to base our selections on those ratings.

In addition to those five *surgery pictures*, we selected five *neutral pictures* from the NAPS catalog, all with consistent ratings of 1 on the disgust scale. These photos include, e.g., a surfer riding a wave and a man walking on the beach with his son. The picture resolution is 1600×1200 for NAPS and 1024×768 for IAPS. For consistency, we rescaled all NAPS pictures to 1024×768 . Thus, the effect of spatial filters will be slightly stronger than for the interview sessions, which used 1600×1200 pictures.

5.2. Techniques

We first selected the six most usable techniques according to surgeons (see Figure 3). We observed that KUWAHARA and BILATERAL yielded almost identical results on our experimental stimuli, and were therefore likely to elicit the same response. Thus, in order to save experimental conditions, we decided to remove BILATERAL, since it is already used as a building block of FLOWABS. We further decided to add HUESHIFT in order to cover a broader spectrum of approaches, even though it was ranked poorly. Looking back at surgeon pictures, we realized that HUESHIFT often gives skin and blood the same green tones. This might be the reason why surgeons found that HUESHIFT suppressed important information. Thus, we changed the shift angle from 120° to -120° in the HSL space, which maps blood and skin to blue/purple tones where color discrimination is superior [BDAB12]. We refer to this technique as HUESHIFT2. Finally, by examining results on experimental stimuli, we realized that the higher number of iterations used for SHAPESIMPL eliminated significantly more details than other techniques. We therefore tuned the settings in order to get more comparable levels of abstraction. We refer to this technique as SHAPESIMPL2. The changes made

[‡] Usage restrictions do not allow us to reproduce or distribute the pictures. These can be obtained from the Nencki Institute of Experimental Biology and CSEA Media.

Table 2: Techniques tested in the experiment.

Abbreviation	Parameter settings
APPARENTGREY	Same as Table 1
HUESHIFT2	$\alpha = -120.0$
SHAPESIMPL2	Same as Table 1 except $N = 4, k = 4$
COHERENCEENH	Same as Table 1
KUWAHARA	Same as Table 1
FLOWABS	Same as Table 1

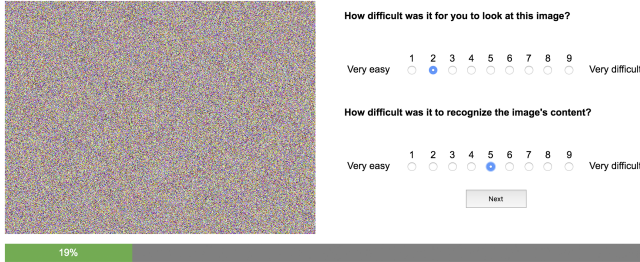


Figure 4: Experiment screen after stimulus presentation.

to SHAPESIMPL could only improve information preservation and the changes made to HUESHIFT do not affect how the technique essentially work (but simply test whether another color could be better). Consequently, the new techniques were not evaluated or validated by surgeons in a new set of interviews. This leaves us with 6 techniques, summarized in Table 2.

5.3. Metrics

A variety of psycho-physiological measures exist to quantify emotional response (see Section 2.3), but they tend to be noisy and they require specialized equipment. We thus decided to simply measure self-reported subjective experience. Psychology has adopted standardized scales to assess affect, such as valence, arousal and disgust scales [LBC97, MZJG14]. However, the difficulty of looking at a repellent surgery image may not directly map to either valence, arousal or disgust as they are understood by participants. Therefore, we chose to ask a more direct question, i.e., “how difficult was it for you to look at this image?”, on a 9-point scale from *very easy* to *very difficult* (see Figure 4). The question was framed in the past tense because, as we will see later on, participants did not see the stimulus when they answered the question.

As a complement to the expert interview, we also asked participants to estimate to what extent the content of the scene has been obfuscated by the filter. The question was “how difficult was it to recognize the image’s content?”, again on a 9-point scale from *very easy* to *very difficult*. The meaning of both questions was explained in preliminary instructions, with examples.

5.4. Design and Procedure

Because between-subject designs typically suffer from low statistical power [BBK14], we used a within-subject design. Each participant saw all combinations of *picture* and *technique*. With 10 pictures (5 surgery, 5 neutral) and 7 techniques (6 + unfiltered), a total of 70 *stimuli* were presented to each participant.

We expected strong ordering effects, as a participant is likely to become less sensitive to surgery pictures after repeated exposure. Furthermore, a participant is more likely to recognize the content of a picture if it had been presented before unfiltered or with a weaker filter. We addressed this in four ways:

- We fully randomized the order of stimulus presentation across participants, with the constraint that two presentations of the same picture had to be separated by at least two other pictures.
- Our experiment instructions warned participants that they would see the same picture multiple times, but asked them to try to answer questions as if they saw each picture for the first time.
- We presented each stimulus for two seconds only, after which we displayed a mask and we invited the participant to answer the two questions (see Figure 4). Limiting exposure to each picture was expected to slow down habituation. At the same time, informal tests we conducted confirmed that a two-second exposure was more than enough to be able to fully scan and recognize a picture.
- We interleaved surgery and neutral pictures, which we expected to further slow down habituation and incite participants to stay focused. Due to the random sequence of surgery and neutral pictures participants could not predict the next image type.

We also used the responses to neutral images to look for participants with poor data. With Likert items, a common and damaging mistake is inversion (e. g., participants replying “easy” instead of “hard” by mistake [SL11, vSSC13]) and we explicitly warned against this mistake in the instructions. Still, we decided (before collecting data) (i) to consider a response of 5 or more to the first question for a neutral image as an obvious inversion, and (ii) to discard the data from all participants who made two or more obvious inversions.

The experiment unfolded as follows: first, participants were given an information sheet and a consent form to sign. We explained that the purpose of the experiment was “to understand people’s levels of aversion in response to processed and unprocessed surgery images”. We also stated that the participant “will be asked to review a series of potentially aversive stimuli and report [their] reactions. The stimuli will include a mix of neutral photos and photos of surgical procedures.”[§] Thus, we tried to be as neutral as possible, and we did not explicitly state that the purpose of the study was to see whether applying filters to surgical images can reduce negative affect. The information sheet provided an example of a surgery image (not used in the study), and asked participants not to participate if they thought they were hypersensitive. It also informed them that they would be free to stop at any time, should they feel too uncomfortable. Then, participants read instructions on a computer (MacBook Pro 2015 with a 2880 × 1800 retina display and a mouse) and completed the 70 trials. Finally, they were given a research debriefing sheet, a brief questionnaire, and were invited to comment on the experiment. The entire experiment lasted between 15 and 20 minutes.

[§] The information sheet and all instructions are available in the online supplementary material (<https://osf.io/4pfes/>).

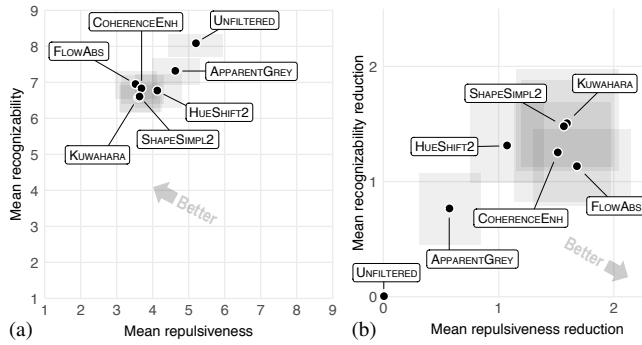


Figure 5: Mean repulsiveness and recognizability ratings for each technique (left); Mean within-subject reduction in repulsiveness and recognizability (right). Boxes are 95% CIs.

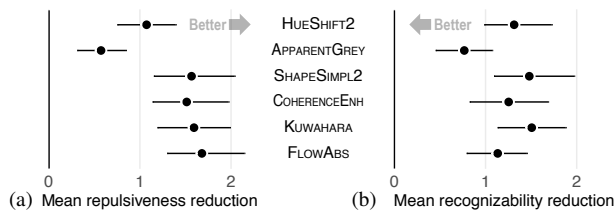


Figure 6: Mean within-subject reduction in repulsiveness and recognizability. Error bars are 95% CIs.

5.5. Participants

We recruited 30 unpaid participants (9 females, age 21–49, mean = 29, med = 26, SD = 8.6), in conformity with our planned sample size. One additional participant made two obvious inversions as defined in Section 5.4 and was therefore discarded from the analysis. Participants were recruited by email posting to our institution and to students we give classes to, as well as word of mouth to neighbor institutions.[¶] Four participants were left-handed and one was ambidextrous. All had normal or corrected-to-normal vision, none were colorblind and they all had higher-education diplomas. Twenty-four reported seeing surgery images or videos before (TV, internet, books), including real surgeries or traumatic injuries ($\times 3$). Eighteen participants participated in a perception study before. Reported sensitivity to surgery images was on average 4.1 (SD = 2.1) on a scale from 1 (not sensitive at all) to 9 (extremely sensitive).

5.6. Quantitative Results

We report and interpret all our results using interval estimation instead of p -values [Cum14, Dra16]. The complete analysis was planned before collecting the data and was preregistered [CGD18] with the Open Science Framework (<https://osf.io/34vzj>).

We report on two dependent variables: *repulsiveness*, which is the response to the first question in Figure 4, and *recognizability*, which is the complement ($y = 10 - x$) of the response to the second

question. For each participant and technique, we averaged responses across all 5 surgery pictures (neutral pictures were not analyzed). Then, for each technique, we derived a point estimate using the mean response across participants, and an interval estimate using the 95% BCa bootstrap confidence interval [KG13].

Figure 5(a) shows point estimates (dots) and interval estimates (gray boxes) within the full space of possible responses. Roughly speaking, each interval indicates the range of plausible values for the population mean, with the point estimate being about seven times more plausible than the interval endpoints [Cum13]. The mean response for unfiltered surgery images is about 5 on the repulsiveness scale and 8 on the recognizability scale. There is strong evidence that all 6 techniques yield smaller average recognizability as well as smaller repulsiveness, except for APPARENTGREY and HUESHIFT2. Unsurprisingly, repulsiveness tends to correlate with recognizability. An ideal processing technique would be an outlier located on the top left of the regression line, but the figure provides no conclusive evidence for such an outlier.

The results of our within-subject study are best examined via average within-subject reduction in repulsiveness and in recognizability [Cum13], summarized in Figure 5(b). We show the same data in Figure 6, with a separate plot for each dependent variable. There is overwhelming evidence that all techniques have a within-subject effect on repulsiveness and on recognizability, on average. The overlap between error bars [KA13] for repulsiveness reduction suggests that APPARENTGREY has the weakest effect, followed by HUESHIFT2, followed by the remaining four. The results for recognizability reduction are less clear, but it is likely that APPARENTGREY yields more recognizability than the other techniques.

Overall, all six techniques are effective, but color manipulation appears to be outperformed by space-domain filtering (SHAPESIMPL2, COHERENCEENH, KUWAHARA, FLOWABS). HUESHIFT2, in particular, is not as effective at making the surgery pictures easier to look at, despite being comparable in terms of preserving informational content. APPARENTGREY, on the other hand, simply appears to be a weaker filter: while it does not reduce repulsiveness dramatically, it better preserves image legibility. Among all techniques, FLOWABS and COHERENCEENH offer the best trade-offs if we only consider point estimates, but given the large overlaps in interval estimates the evidence is very weak.

5.7. Participant Feedback

Among the ones who offered feedback, two found FLOWABS to be the best technique. Several participants mentioned that it looked like a cartoon or comic-strip ($\times 8$) making surgery photos easier to look at ($\times 8$). A single participant found that it could make it harder because of the more salient contours, while three participants said that it makes content easier to recognize. HUESHIFT2 was said to make photos easier to look at ($\times 3$) but the content harder to recognize ($\times 3$). Four participants found that it made content appear “unnatural” or “disturbing”, while one mentioned that it makes patients look like aliens. The APPARENTGREY was deemed easier to look at ($\times 3$) and easier to recognize ($\times 1$), while two participants reported it was harder to recognize the content without colors. Additionally, one participant thought that the APPARENTGREY’s effectiveness

[¶] Although our experiment was designed to be runnable on crowdsourcing platforms, we only got IRB approval for running lab studies.

depends on the input picture. This was also reflected by another participant who mentioned that the best technique could be picture-dependent. COHERENCEENH, KUWAHARA, and SHAPESIMPL2 were reported to be the best technique by one participant each, and to make content harder to recognize by three participants.

5.8. Summary of Findings

We found that all six tested techniques make it easier to look at surgery photos as judged by lay people, even if the effects are moderate and come at a slight cost in legibility. Combining quantitative and qualitative results from both studies, two particularly successful techniques are FLOWABS (structure-adaptive filtering [KD08]) and COHERENCEENH (coherence-enhancing filtering [KK11]). FLOWABS is the only technique integrating edge enhancement, which seems to provide a double benefit: it may increase legibility and recognizability by emphasizing contours, while also producing a cartoon appearance that may make photos look less “real” and thus less disturbing. While this highlights the potential of edge-enhancing techniques, our results still show that other approaches can also be very successful and should still be considered. Concerning COHERENCEENH and the almost equally successful KUWAHARA, SHAPESIMPL, and BILATERAL, they all provide some form of nonlinear/coherence-enhancing diffusion [Wei99] or anisotropic filtering in order to obtain a kind of a painterly look, although not as explicit as OILPAINT. This subtle painterly look may also make the photos look less real, without compromising too much in legibility. COHERENCEENH was found to be the most legible by surgeons, perhaps because compared to SHAPESIMPL, for example, it preserves “the shape by using a curvature preserving smoothing method that enhances coherence” [KK11] and is more resistant to noise.

The finding that color manipulation was less effective came to us as a surprise, given the emotional force blood appears to carry, and the tradition of altering its color in video games and Japanese animation. Not only HUESHIFT was found to discard information despite the absence of spatial filtering, it was also found to be weaker at affective dampening compared to the approaches mentioned before. APPARENTGREY may still have useful applications as a weak surgery filter, as it was found to be relatively usable by surgeons.

Almost all artistic stylization techniques (WATERCOLOR, BRUSHSTROKES, HATCHING and STIPPLING) were found by surgeons to discard too much information to be useful. However, OILPAINT was relatively well received, even if it did not make it in our final experiment. This, together with the fact that FLOWABS is the filter that comes closest to stylization, suggest that stylization still holds promise and deserves to be investigated further.

6. Second Interview with Surgeons

Our experiment with lay people confirmed that processing surgical photos can diminish negative affect. In principle, it is thus possible to provide a tool that automatically processes surgical photos to reduce affective responses, without losing too much information. Our study suggests that FLOWABS, COHERENCEENH, and KUWAHARA would be good candidates for filters to support in such a tool. Yet it is still unclear what would be good default filter settings and whether our approach extends to video processing. We thus conducted a second set of interviews with surgeons.

6.1. Abstraction Parameters

We used the parameters summarized in Section 3.5 as a starting point to derive a broader range of parameter settings to use in our interview with surgeons. To simplify the surgeons’ tasks, we wanted them to adjust a single “level of abstraction” parameter instead of multiple technical parameters. We thus had to determine the parameters of each technique (see Table 1) that subjectively influence the overall level of abstraction of the output and map well-defined configurations to this level of abstraction parameter.

FlowAbs. FLOWABS superimposes the results of an orientation-aligned bilateral filtering and DoG filtering to obtain cartoon-like effects, thus they need to be independently controlled from each other. Because FLOWABS involves iterative color smoothing, we decided to parameterize the number of iterations of bilateral filtering, i. e., n_a used for color abstraction and n_e for DoG filtering [KD08]. We favored this approach over increasing the range weights to avoid blurred edges in high levels of abstraction [TM98]. Because bilateral filtering is used as pre-processing for edge enhancement, i. e., for denoising, we also had to parameterize the DoG filter to explicitly control the amount of edges. Edges are detected at the zero-crossings of the DoG, whose amplitude can be adjusted by varying the strength of the inhibitory effect of the larger Gaussian, controlled by parameter τ [WKO12], to obtain more or fewer contour enhancements. Note that Kyprianidis et al. [KD08] originally defined $n_e < n_a$ as a constraint for pre-smoothing, which we removed to also obtain results that combine deliberate contour enhancements with subtle color abstraction. The abstraction defined by FLOWABS is thus independently defined by the parameter n_a for the level of abstraction in color space and (n_e, τ) for contour enhancement.

CoherenceEnh. As proposed by Kyprianidis and Kang [KK11], we chose to adjust the number of iterations N to control the amount of adaptive smoothing and shock filtering. By maintaining stable parameters for each iteration of COHERENCEENH, we thus achieve a good balance between enhancing directional image features and smoothing of isotropic image regions. This approach is also able to preserve singular image features at a reasonable scale while providing easy control over the level of abstraction.

Kuwahara. The KUWAHARA filter proposed by Kyprianidis [Kyp11] uses a multi-scale computation scheme that propagates local orientation estimates and filtering results up a low-pass filtered image pyramid. To this end, adjusting the level of abstraction can be easily controlled by defining up to which level of the image pyramid (from coarse to fine order) the response of the single-scale anisotropic Kuwahara filter should be merged, i. e., for which we adjust parameter p_s [Kyp11]. This approach enables us to achieve high levels of abstraction of large homogeneous color regions without noticeable artifacts in large low-contrast regions.

For each technique and effect, we decided to define 10 levels of abstraction and sensitivity of contour enhancement that range from photorealism, i. e., the input image is used as output, to strong filtering effects according to the limiting cases reported by the respective works. We provide an overview of these configurations in Table 3.

Table 3: Configurations used to provide different levels of abstraction (LoA) and sensitivities of contour enhancement (CE) for the second interview. The parameters in Table 1 are used as a basis.

LoA	FLOWABS	KUWAHARA	COHERENCEENH	CE	FLOWABS
0	Original	Original	Original	0	$n_e = 20, \tau = 0.0$
1	$n_d = 0$	$p_s = 1.0$	$N = 1$	1	$n_e = 20, \tau = 0.93$
2	$n_d = 1$	$p_s = 0.7$	$N = 2$	2	$n_e = 16, \tau = 0.94$
3	$n_d = 3$	$p_s = 0.5$	$N = 4$	3	$n_e = 13, \tau = 0.95$
4	$n_d = 5$	$p_s = 0.4$	$N = 8$	4	$n_e = 10, \tau = 0.96$
5	$n_d = 7$	$p_s = 0.35$	$N = 12$	5	$n_e = 7, \tau = 0.97$
6	$n_d = 10$	$p_s = 0.3$	$N = 20$	6	$n_e = 5, \tau = 0.98$
7	$n_d = 13$	$p_s = 0.25$	$N = 30$	7	$n_e = 3, \tau = 0.99$
8	$n_d = 16$	$p_s = 0.2$	$N = 40$	8	$n_e = 1, \tau = 0.99$
9	$n_d = 20$	$p_s = 0.15$	$N = 50$	9	$n_e = 0, \tau = 0.99$

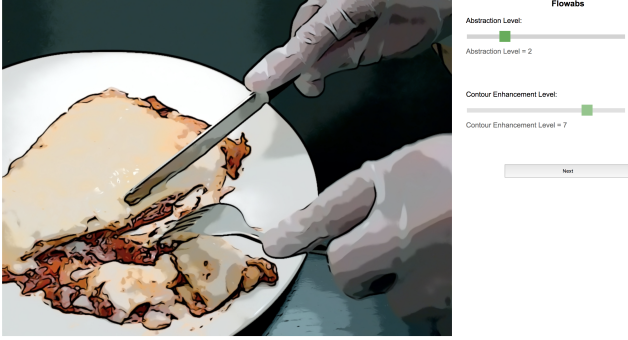


Figure 7: Interface used by surgeons to determine good defaults.

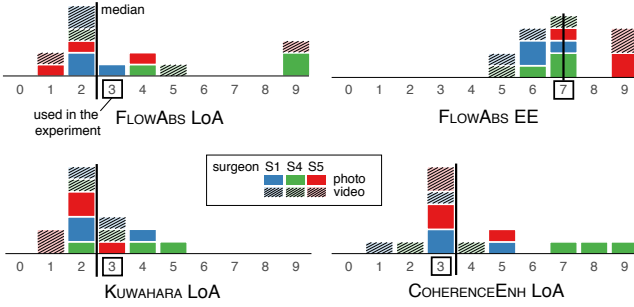


Figure 8: Parameter settings used by the surgeons. Each square depicts the setting used in a specific photo (uniform fill) or video (hatched fill) by one of the three surgeons (red, blue, or green). The vertical line is the median setting. The setting used in the previous experiment is shown with a black square. See Table 3 for the correspondence between the numbers 0–9 and parameter values.

6.2. Participants

Two of the surgeons we previously interviewed, *S1* and *S4* (see Section 4), agreed to also participate in this second interview. We also recruited an additional orthopedic surgeon *S5* (15 years experience). While *S5* was not involved in our research at the time of this second interview, *S1* and *S4* were already co-authors of the previously published conference paper [BSB*18].

Similar to the first batch of interviews, we asked each surgeon to send us three of their own surgery photos (for *S1* and *S4* we used the ones they sent us before) and two videos of their own surgical interventions. We cropped (when necessary) and resized them to 1024×768 , before processing them as detailed in Table 3. For

each photo and video, we thus obtained $10 \times 10 = 100$ stimuli for FLOWABS, and 10 stimuli for COHERENCEENH and KUWAHARA.

The interviewer first summarized the results from the first interview and the experiment on lay people, and then explained the goal of the current interview. We then presented the surgeons with an interface that lets them change filter settings on their own photos and videos (Figure 7), using a 15-inch MacBook Pro (2880×1800 retina display) with a computer mouse. We instructed them as follows: “Suppose you want to use this image/video to explain a procedure to a patient/student/journalist. I would like you to show me the maximum level of abstraction you think you can use. That is, we are trying to determine the level that, if exceeded, will mask the important information in the photo.” The interviewer then explained that the more the cursor is moved to the right (see Figure 7), the more abstraction is introduced. We encouraged the surgeons to think aloud.

We first presented their three photos to the surgeons (Part 1), and then their two videos (Part 2). For each photo/video, they had to set the cursor(s) three times in a row, once per technique. We counterbalanced the order in which surgeons saw the techniques. After they were done with a photo/video, surgeons saw a screen where the three processed photos/videos were displayed again, and were asked to state which technique was their favorite and why. At the end of the interview, we asked the surgeons for additional comments. Each interview lasted approximately one hour.

6.3. Choices of Settings

We summarize the settings chosen by surgeons in Figure 8. While there is lots of variation across photos/videos and surgeons, there are also clear trends. Values for FLOWABS’s contour enhancing (CE) that are less than 5 were never used. Also, the chosen abstraction levels remain low for all techniques: the median value is close to 2–3, and values greater than 4 are either never used (KUWAHARA) or rarely used (COHERENCEENH, FLOWABS). We discuss the reasons for this choice in the next section on qualitative feedback.

We can also note that, overall, the settings used for videos do not differ noticeably from the settings used for photos, suggesting that the same defaults can be used in an automated tool, irrespective of whether the content is a photo or a video. In addition, for all four parameters, the median value is remarkably close to the settings we used in our experiment with lay people (see Section 5). Since the settings used in our experiment (and which were shown to reduce negative affect) are similar to the settings chosen by surgeons (and which were judged to preserve important information), we can be reasonably confident that it is possible to tune filters so that they simultaneously preserve information and reduce negative affect.

Based on the median responses (vertical lines in Figure 8), we recommend the following default settings for 1024×768 surgery photos and videos: FLOWABS: $n_d = 2$; $n_e = 3$; $\tau = 0.99$, KUWAHARA: $p_s = 0.6$, COHERENCEENH: $N = 6$. For the cases where the median is a non-integer, we derived these values by linear interpolation. For other parameter values, please refer to Table 1.

Note that the default settings values we derived from our experiment are only central tendency estimates. Given the large variability in the settings used (see Figure 8), they are not expected to work

universally. More than half of the time, better filtered images will be produced if a surgeon is allowed to fine tune the settings depending on the nature of the photo or video. Furthermore, the central tendency estimates were computed from a sample of 3 participants, so they are very imprecise. Nevertheless, these values provide a good empirically-based starting point for choosing default settings values.

As mentioned in Section 3.5, parameters of the techniques that relate to spatial distance can be scaled to obtain a stable level of abstraction for images and videos with varying resolution. This refers to $\{\sigma_d, \sigma_m\}$ for FLOWABS, $\{\sigma_d, \sigma_t, \sigma_g, \sigma_a\}$ for COHERENCEENH, and $\{\rho, p_s\}$ for KUWAHARA (see Table 1).

6.4. Qualitative Feedback and Preferences

Of all images and videos that we showed to surgeons at the end of the interview with their personal maximum degree of abstraction, FLOWABS was preferred nine times, COHERENCEENH four times, and KUWAHARA twice. Interestingly, two surgeons (*S4* and *S5*) commented that their preferences depend on the context: they explained they would chose one technique for its artistic/aesthetic aspect, but another for its efficiency to convey information. *S5*, for instance, reported that he preferred FLOWABS for all images for communication purposes, but that COHERENCEENH ($\times 1$ image) and KUWAHARA ($\times 2$) were more aesthetically pleasing. Similarly, *S4* reported for her third surgery picture to prefer showing FLOWABS to patients, but preferred the artistic aspect of KUWAHARA.

We strongly encouraged surgeons to voice comments throughout the interview. In general, and similar to our first interview with them and its results (Figure 3), they liked all three techniques and positively commented on the aesthetic aspects they can provide to the images or videos. We now report the most interesting and recurrent comments. First, while all thought that our approach was promising for pictures, they were positively surprised and impressed by the quality of the videos. One surgeon (*S4*) even commented on a FLOWABS-processed video “oh, it is sick.” The cartoon/anime/drawing effect of FLOWABS was again mentioned during the interviews (3 \times in total). The aesthetics of KUWAHARA was also praised: two surgeons mentioned that it was beautiful or nice. It was also reported twice that KUWAHARA gave a pastel-like rendering to the stimuli. *S1* also mentioned the aesthetics of COHERENCEENH, saying that it can also give a cartoon effect.

It is worth noting that one of the surgeons (*S1*) particularly highlighted that the end result with a single technique depends on the input images/videos by comparing multiple times the results he obtained with the same technique on different images or videos.

For FLOWABS, *S5* mentioned that, for a “didactic/pedagogical purpose, it seems that always having the highest level of contour enhancing is best,” and on videos the “highest levels for both [contour enhancement and abstraction make the video] really look [...] like an anime.” He also stated that KUWAHARA “is very nice for reducing the emotional impact of videos, it would be interesting to combine this with the contour enhancement feature [from FLOWABS].” Finally, *S5* also suggested to combine multiple techniques for different parts of an image. For instance, as KUWAHARA was often praised by surgeons for its aesthetics, it would be interesting to use it for an image background, while the foreground and details could be

processed with FLOWABS. This could yield a “pretty” background/context with a clear and more contoured focus. *S5* also suggested that he would be interested in having an interactive tool that allows him to specifically select where the contours should be enhanced: he wants to start from a normal image and iteratively highlight contours, before blurring out all unimportant details. He also stated that it would be interesting to have the possibility to simplify the contours’ shape so that they look more like familiar shapes to patients.

6.5. Summary of Findings

This second batch of interviews with three surgeons confirmed that the techniques highlighted by our study with lay people were indeed of interest to all three surgeons. We also found that they seem to be efficient when applied to videos, not only to still images. We believe that the median parameter values that we have gathered through these interviews can be used as a defaults when processing surgery images. The interviews also confirmed that FLOWABS’s contour enhancement was of particular interest: it highlights and helps to distinguish important parts of an image, while it also accentuates the “cartoon” effect which has been previously mentioned by both lay people and surgeons to reduce the revolting aspects of surgery. We also found that, if such contour enhancement is possible, having a high contour-enhancement value has been considered, for both videos and images, to improve the legibility of the image or video.

7. Arkangel: A Chrome Extension

The techniques and results presented here could be used directly in the creation of content from surgeons to assist them during procedure explanations to patients or medical vulgarization. However, we also present ARKANGEL^{||}, a Google chrome extension that assists internet users in getting access to information, while reducing the impact of informative but potentially shocking visual content.

Google and Wikipedia, e.g., are used for everyday life research [HE09] and, in particular, health information [CWS*17, GKN*05, KH18, RMSS15, SWR08, Yan10]. A young audience often uses them to inform themselves [FCBS18, SBPG03, GKN*05, SWR08] or work on school projects [BCAL15, She18]—yet they can contain graphic or inappropriate content [LSG*18, LKPS14, Wik03, Wik10b]. Interestingly, social media sites (e.g., Facebook, Twitter) currently allow people to publish shocking content [Fac18, Gil18, Twi18]. Similarly, Wikipedia contains images or videos that can be judged as shocking or inappropriate to some audience [Wik03, Wik10a, Wik10b]. News media are also not exempt from the use of graphic media to report on the consequences of war [Eli14] or the reality of carnage [Too14]. Some even argue that their use is now more frequent than ever before [Pos14]. While removing graphic content is possible, it is often argued that graphic content informs the public and raises awareness [Fac18, PET08, Wik10a] and should, consequently, remain available. Moreover, studies have shown that blocking software intended to protect children also blocks access to useful information [RRH*02]. Facebook and Twitter thus

^{||} <https://github.com/lonnibesancon/Arkangel>

implemented a warning screen [Fac18, Twi18] and Wikipedia is interested in allowing users to customize their viewing experience of the encyclopedia [Wik10a].

A reduction of the affective responses to online pictures/videos could thus yield advantages for public learning in the context of medical/surgical information and online media. We demonstrated that FLOWABS can reduce the affective response to such media, and we thus implemented the technique in ARKANGEL as a possible solution for such contexts. ARKANGEL dynamically parses HTML images, initially obfuscates their contents via heavy Gaussian blurring, and then replaces the obfuscated contents by the processed versions using a WebGL implementation of FLOWABS. For image parsing, ARKANGEL uses a *MutationObserver* of the JavaScript API that listens on changes in the DOM while browsing web pages, and WebGL Canvas as an interface between original unprocessed and processed image data using data URIs for embedding. For processing, we use the values gathered in our second batch of interviews as default values. All images are automatically processed and we added the possibility to change the level of abstraction and sensitivity of contour enhancement via extension options, and revert to the original images via context menus. Ultimately, an improved version of our extension would automatically detect content that is likely to produce strong affective responses through image processing (such as already existing approaches to detect blood [Goo] or pornography [Goo, MS16, PSL14, ZFBL17], leveraging the power of machine learning) or code-based analysis to detect pre-tagged images (such as Facebook's tagged graphic images or videos [Fac18]).

8. Discussion and Future Work

In this section, we put our results into context, mention limitations of our study, and conclude with a discussion of future work.

8.1. Summary of Contributions

Our main contribution is that the six techniques we included in our experiment with lay people all make it easier to look at surgery photos, even if the effects are moderate and come at a slight cost in legibility. Combining quantitative and qualitative results from the study and interviews, two particularly successful techniques are FLOWABS (structure-adaptive filtering [KD08]) and COHERENCEENH (coherence-enhancing filtering [KK11]). FLOWABS is the only technique that integrates contour emphasis, which seems to provide two benefits: it may increase legibility and recognizability by emphasizing contours, while also producing a cartoon appearance that may make photos look less “real” and, thus, less disturbing. Concerning COHERENCEENH and the almost equally successful KUWAHARA, SHAPESIMPL, and BILATERAL, they all provide some form of nonlinear/coherence-enhancing diffusion [Wei99] or anisotropic filtering and create a somewhat painterly look, although not as explicit as OILPAINT. This subtle painterly look may also make photos look less real, without compromising legibility too much. We found COHERENCEENH to be the most legible by surgeons, perhaps because—compared to, e.g., SHAPESIMPL—it preserves “the shape by using a curvature preserving smoothing method that enhances coherence” [KK11] and is more robust to noise.

The finding that color manipulation was less effective came to

us as a surprise, given the emotional force blood appears to carry and the tradition of altering its color in video games and Japanese animation. Not only did we find HUESHIFT to discard information despite the absence of spatial filtering, we also found to be weaker at affective dampening compared to the approaches mentioned before. APPARENTGREY may still have useful applications as a weak surgery filter, as we found it to be relatively usable by surgeons.

We found almost all artistic stylization techniques (WATERCOLOR, BRUSHSTROKES, HATCHING, and STIPPLING) to discard too much information to be useful. OILPAINT, however, was relatively well received, even if it did not make it in our experiment with lay people. Together with the fact that FLOWABS is the filter that comes closest to stylization, these observations suggest that stylization holds promise and deserves to be investigated further.

We found the three techniques we included in our second set of interviews (FLOWABS, COHERENCEENH, and BILATERAL) to be usable on both videos and pictures. We also derived default parameters for these techniques that preserve enough information so that surgeons can still use them to explain procedures to lay people.

The discussed overwhelming presence of graphic content in the news and social media is problematic, in particular for young or sensitive audiences. Some of these media possess similar properties with surgical images (e.g., blood or visible organs). We thus believe that FLOWABS, COHERENCEENH, and BILATERAL can reduce the affective impact of such media, thus making ARKANGEL a useful tool for such audiences. Their effectiveness on other forms of graphic content (e.g., nudity or pornography) is currently speculative, but we hope that our initial investigation will spark follow-up experiments in the graphics or psychology communities.

8.2. Limitations and Future Work

This article is an initial investigation but opens up exciting avenues for future research. We discuss the limitations of our interviews, of the experiment with lay people, and of the Google Chrome extension as well as directions for future work next.

Concerning the first set of interviews, we only examined 13 techniques and many others remain to be tested (e.g., [KD08]). More sophisticated (and possibly more effective) variants exist, especially in the area of color manipulation. Combinations of techniques also merit consideration, such as using contour enhancement from FLOWABS with other techniques, or combining color manipulation with spatial filters. As future work, we thus also plan to investigate how neural style transfer [GEB16, JAFF16, SID17] can be used to replicate illustrations found in medical textbooks, which could be extended by a loss function that weights disgust or takes into account the target group to auto-adjust the abstraction. Such techniques can be quite promising in our context, although their outcome is still hard to control, in particular in an illustrative context [SID17]. We also consider involving surgeons with other specializations to investigate whether it has an impact on which technique to use.

Similarly, we have only considered a limited number of input images and videos from the surgeons for the second batch of interviews (3 images & 2 videos, per surgeon). Such a reduced set was used to limit the interview duration. While the processing techniques

we used were, once again, evaluated to be efficient, the end results highly depends on the input image. As a consequence, the median levels we have identified through our interviews might not always be ideal for all surgery images/videos. Future work in this direction could rely on ARKANGEL to collect data from users browsing shocking content for informative purposes such as new medical apprentices who might not be accustomed yet to surgical media.

Our experiment with lay people is limited by our self-reported measures that may not truly reflect the participants' experiences. In particular, we cannot rule out the possibility that responses were tainted by a social desirability bias [Fis93] or, worse, by a good-subject effect [NM08]. This could be improved in the future by adopting between-subject designs and/or objective measurements (Section 2.3), but it is unclear whether such methods would provide sufficient sensitivity for populations that are not hypersensitive or BII-phobic. Second, our results only apply to the specific surgery pictures we tested. To draw general inferences about the effectiveness of the techniques, one would need to design a study where stimuli are randomly drawn from a large collection of representative surgery photos [JWK12]. A third important limitation is that we only compared techniques with specific settings, so our findings do not necessarily apply to the techniques in general. Solutions to this issue include testing many settings or devising a systematic calibration that can determine parameters for fair comparisons.

A pending challenge is the use of a standardized picture set for future studies. Even though developments are underway to specify benchmark pictures for expressive rendering [MR16, MR17, RMB*17], we had to resort to the NAPS and IAPS catalogs, yet even in these sources only few pictures met our requirements. While the selected pictures are available for future work through the named sources, it may be useful to specify a set of standard pictures to study the impact of stylization on affect.

As we noted in Section 7, only part of the online shocking content is based on surgical procedures. We only empirically demonstrated the effectiveness of FLOWABS on surgery content, and it remains unclear how it would fare for other shocking content such as other type of medical images (e.g., skin diseases), as well as user-generated content featuring nudity and pornography, injuries, or violence. These can psychologically impact professionals who report [Joh18, Juk17, SDN18] or monitor it [FAW14]. In particular, pornography is often found inadvertently by users and especially young audiences [Hal14, LSG*18, MMA*16, SP14, WTSHS15] while browsing for information. In addition to society's willingness to protect young users against this content [Mad10] that could lead to behavioral modifications [OBMR12, PV09, PV16, WTK15] through content blocking [CAS*16, MB10, PAM*17, PE10], children also report to be worried about these inappropriate media [Hal14, LKPS14, TBO*17]. While we cannot predict how FLOWABS and ARKANGEL could help with the issue of nudity and pornography, it is interesting to notice that violent or accident media contain features that are common to the surgery images and videos we have studied here: they include injuries, blood, and visible internal organs. We believe it is thus reasonable to assume that FLOWABS and ARKANGEL can also reduce the affective responses of media displaying these features and could therefore be used to browse websites covering these topics, be it user-generated content

on social platforms or news media. Nonetheless, further work is needed to assess whether FLOWABS works as well in these contexts or whether other filters could yield better results.

Acknowledgements

Thanks to Steve Haroz, Yvonne Jansen, Wesley Willett, Pascal Guitton, and Heiko Thiel for their help and feedback. This research was partly funded by the Federal Ministry of Education and Research (BMBF), Germany, for the AVA project 01IS15041B.

References

- [AHO13] ARMSTRONG T., HEMMINGER A., OLATUNJI B. O.: Attentional bias in injection phobia: Overt components, time course, and relation to behavior. *Behaviour Research and Therapy* 51, 6 (June 2013), 266–273. doi: 10.1016/j.brat.2013.02.008 2
- [BBK14] BELLEMARE C., BISSONNETTE L., KRÖGER S.: *Statistical power of within and between-subjects designs in economic experiments*. IZA Discussion Papers, No. 8583, IZA Institute of Labor Economics, Oct. 2014. URL: <https://www.iza.org/pub/z2nqxLYL>. 8
- [BCAL15] BEHESHTI J., COLE C., ABUHIMED D., LAMOUREUX I.: Tracking middle school students' information behavior via Kuhlthau's ISP model: temporality. *Journal of the Association for Information Science and Technology* 66, 5 (May 2015), 943–960. doi: 10.1002/asi.23230 12
- [BDAB12] BACHY R., DIAS J., ALLEYSSON D., BONNARDEL V.: Hue discrimination, unique hues and naming. *Journal of the Optical Society of America A* 29, 2 (Feb. 2012), A60–A68. doi: 10.1364/JOSA.29.000A60 7
- [BKTS06] BOUSSEAU A., KAPLAN M., THOLLOT J., SILLION F. X.: Interactive watercolor rendering with temporal coherence and abstraction. In *Proc. NPAR* (2006), ACM, New York, pp. 141–149. doi: 10.1145/1124728.1124751 4, 5, 6
- [BLD*08] BENUZZI F., LUI F., DUZZI D., NICHELLI P. F., PORRO C. A.: Does it look painful or disgusting? Ask your parietal and cingulate cortex. *Journal of Neuroscience* 28, 4 (Jan. 2008), 923–931. doi: 10.1523/JNEUROSCI.4012-07.2008 2
- [BSB*18] BESANÇON L., SEMMO A., BIAU D., FRACHET B., PINEAU V., SARIALI E. H., TAOUACHI R., ISENBERG T., DRAGICEVIC P.: Reducing affective responses to surgical images through color manipulation and stylization. In *Proc. Expressive* (2018), ACM, New York, pp. 4:1–4:13. doi: 10.1145/3229147.3229158 1, 11
- [CA12] CHAPMAN H. A., ANDERSON A. K.: Understanding disgust. *Annals of the New York Academy of Sciences* 1251, 1 (Mar. 2012), 62–76. doi: 10.1111/j.1749-6632.2011.06369.x 2
- [Čad08] ČADÍK M.: Perceptual evaluation of color-to-grayscale image conversions. *Computer Graphics Forum* 27, 7 (Oct. 2008), 1745–1754. doi: 10.1111/j.1467-8659.2008.01319.x 3
- [Can86] CANNY J.: A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 8, 6 (Nov. 1986), 679–698. doi: 10.1109/TPAMI.1986.4767851 5
- [CAS*97] CURTIS C. J., ANDERSON S. E., SEIMS J. E., FLEISCHER K. W., SALESIN D. H.: Computer-generated watercolor. In *Proc. SIGGRAPH* (1997), ACM, New York, pp. 421–430. doi: 10.1145/258734.258896 3
- [CAS*16] CAETANO C., AVILA S., SCHWARTZ W. R., GUIMARÃES S. J. F., DE A. ARAÚJO A.: A mid-level video representation based on binary descriptors: A case study for pornography detection. *Neurocomputing* 213 (Nov. 2016), 102–114. doi: 10.1016/j.neucom.2016.03.099 14
- [CCF97] CARPENDALE M. S. T., COWPERTHWAIT D. J., FRACCHIA F. D.: Making distortions comprehensible. In *Proc. VL* (1997), IEEE Computer Society, Los Alamitos, pp. 36–45. doi: 10.1109/VL.1997.626556 3

- [CGD18] COCKBURN A., GUTWIN K., DIX A.: HARK no more: On the preregistration of CHI experiments. In *Proc. CHI* (2018), ACM, New York, pp. 141:1–141:12. doi: 10.1145/3173574.3173715 9
- [CMM02] COMANICIU D., MEER P., MEMBER S.: Mean shift: A robust approach toward feature space analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 24, 5 (May 2002), 603–619. doi: 10.1109/34.1000236 3
- [COL09] CISLER J. M., OLATUNJI B. O., LOHR J. M.: Disgust, fear, and the anxiety disorders: A critical review. *Clinical Psychology Review* 29, 1 (Feb. 2009), 34–46. doi: 10.1016/j.cpr.2008.09.007 2
- [CSESS05] COSTA SOUSA M., EBERT D. S., STREDNEY D., SVAKHINE N. A.: Illustrative visualization for medical training. In *Proc. CAE* (2005), Eurographics Association, Goslar, Germany, pp. 201–208. doi: 10.2312/COMPAESTH/COMPAESTH05/201-208 1
- [CSR10] CRIMINISI A., SHARP T., ROTHER C., PÉREZ P.: Geodesic image and video editing. *ACM Transactions on Graphics* 29, 5 (Oct. 2010), 134:1–134:15. doi: 10.1145/1857907.1857910 5
- [Cum13] CUMMING G.: *Understanding the new statistics: Effect sizes, confidence intervals, and meta-analysis*. Routledge, New York, 2013. doi: 10.4324/9780203807002 9
- [Cum14] CUMMING G.: The new statistics: Why and how. *Psychological Science* 25, 1 (Jan. 2014), 7–29. doi: 10.1177/0956797613504966 9
- [CWS*17] CHU J. T., WANG M. P., SHEN C., VISWANATH K., LAM T. H., CHAN S. S. C.: How, when and why people seek health information online: Qualitative study in Hong Kong. *Interactive Journal of Medical Research* 6, 2 (Dec 2017), e24:1–e24:10. doi: 10.2196/ijmr.7000 12
- [DBHM03] DUKE D. J., BARNARD P. J., HALPER N., MELLIN M.: Rendering and affect. *Computer Graphics Forum* 22, 3 (Sept. 2003), 359–368. doi: 10.1111/1467-8659.00683 1
- [Dra16] DRAGICEVIC P.: Fair statistical communication in HCI. In *Modern Statistical Methods for HCI*, Robertson J., Kaptein M., (Eds.). Springer International Publishing, Cham, Switzerland, 2016, ch. 13, pp. 291–330. doi: 10.1007/978-3-319-26633-6_13 9
- [DS02] DECARLO D., SANTELLA A.: Stylization and abstraction of photographs. *ACM Transactions on Graphics* 21, 3 (July 2002), 769–776. doi: 10.1145/566654.566650 3
- [DWI13] DRAGICEVIC P., WILLETT W., ISENBERG T.: Illustrative data graphic style elements. In *Posters at Expressive* (2013). 1
- [ECS06] EBERT D. S., COSTA SOUSA M. (Eds.): *Illustrative Visualization for Medicine and Science* (2006), vol. 6 of *ACM SIGGRAPH 2006 Course Notes*, ACM. doi: 10.1145/1185657.1185691 1
- [Eli14] ELLIOTT T.: ‘Reality of war’ laid bare in publication of graphic images of MH17, Gaza dead. *The Sydney Morning Herald* (July 2014). URL: <https://www.smh.com.au/world/reality-of-war-laid-bare-in-publication-of-graphic-images-of-mh17-gaza-dead-20140725-zwzow4.html>. 12
- [EVW*07] EIPPERT F., VEIT R., WEISKOPF N., ERB M., BIRBAUMER N., ANDERS S.: Regulation of emotional responses elicited by threat-related stimuli. *Human Brain Mapping* 28, 5 (May 2007), 409–423. doi: 10.1002/hbm.20291 2
- [Fac18] FACEBOOK: Facebook community standards, violence and graphic content, October 2018. URL: https://www.facebook.com/communitystandards/graphic_violence. 12, 13
- [FAW14] FEINSTEIN A., AUDET B., WAKNINE E.: Witnessing images of extreme violence: A psychological study of journalists in the newsroom. *JRSM Open* 5, 8 (Aug. 2014), 2054270414533323. doi: 10.1177/2054270414533323 14
- [FCBS18] FREEMAN J. L., CALDWELL P. H. Y., BENNETT P. A., SCOTT K. M.: How adolescents search for and appraise online health information: A systematic review. *The Journal of Pediatrics* 195 (Apr. 2018), 244–255.e1. doi: 10.1016/j.jpeds.2017.11.031 12
- [FFLS08] FARBMAN Z., FATTAL R., LISCHINSKI D., SZELISKI R.: Edge-preserving decompositions for multi-scale tone and detail manipulation. *ACM Transactions on Graphics* 27, 3 (Aug. 2008), 67:1–67:10. doi: 10.1145/1360612.1360666 5
- [Fis93] FISHER R. J.: Social desirability bias and the validity of indirect questioning. *Journal of Consumer Research* 20, 2 (Sept. 1993), 303–315. doi: 10.1086/209351 14
- [FPC*14] FARIDUL H. S., POULI T., CHAMARET C., STAUDER J., TRÉMEAU A., REINHARD E.: A survey of color mapping and its applications. In *Eurographics State of the Art Reports* (2014), Eurographics Association, Goslar, Germany, pp. 43–67. doi: 10.2312/egst.20141035 3
- [GD07] GRUNDLAND M., DODGSON N. A.: Decolorize: Fast, contrast enhancing, color to grayscale conversion. *Pattern Recognition* 40, 11 (Nov. 2007), 2891–2896. doi: 10.1016/j.patcog.2006.11.003 3
- [GD12] GILCHRIST P. T., DITTO B.: The effects of blood-draw and injection stimuli on the vasovagal response. *Psychophysiology* 49, 6 (June 2012), 815–820. doi: 10.1111/j.1469-8986.2012.01359.x 2
- [GDA*12] GERSTNER T., DECARLO D., ALEXA M., FINKELSTEIN A., GINGOLD Y., NEALEN A.: Pixelated image abstraction. In *Proc. Expressive* (2012), Eurographics Association, Aire-la-Ville, Switzerland, pp. 29–36. doi: 10.2312/PE/NPAR/NPAR12/029-036 3
- [GEB16] GATYS L. A., ECKER A. S., BETHGE M.: Image style transfer using convolutional neural networks. In *Proc. CVPR* (2016), IEEE Computer Society, Los Alamitos, pp. 2414–2423. doi: 10.1109/CVPR.2016.265 13
- [GG01] GOOCH B., GOOCH A. A.: *Non-Photorealistic Rendering*. A K Peters, Ltd., Natick, 2001. doi: 10.1201/9781439864173 1
- [Gil18] GILLESPIE T.: *Custodians of the Internet: Platforms, content moderation, and the hidden decisions that shape social media*. Yale University Press, 2018. 12
- [GKN*05] GRAY N. J., KLEIN J. D., NOYCE P. R., SESSELBERG T. S., CANTRILL J. A.: Health information-seeking behaviour in adolescence: The place of the internet. *Social Science & Medicine* 60, 7 (Apr. 2005), 1467–1478. doi: 10.1016/j.socscimed.2004.08.010 12
- [GLJ*10] GOOCH A. A., LONG J., JI L., ESTEY A., GOOCH B. S.: Viewing progress in non-photorealistic rendering through Heinlein’s lens. In *Proc. NPAR* (2010), ACM, New York, pp. 165–171. doi: 10.1145/1809939.1809959 0
- [Goo] GOOGLE: Cloud vision. URL: <https://cloud.google.com/vision/>. 13
- [GOTG05] GOOCH A. A., OLSEN S. C., TUMBLIN J., GOOCH B.: Color2Gray: Saliency-preserving color removal. *ACM Transactions on Graphics* 24, 3 (July 2005), 634–639. doi: 10.1145/1073204.1073241 3
- [GRG04] GOOCH B., REINHARD E., GOOCH A.: Human facial illustrations: Creation and psychophysical evaluation. *ACM Transactions on Graphics* 23, 1 (Jan. 2004), 27–44. doi: 10.1145/966131.966133 1, 5
- [GW02] GOOCH A. A., WILLEMSSEN P.: Evaluating space perception in NPR immersive environments. In *Proc. NPAR* (2002), ACM, New York, pp. 105–110. doi: 10.1145/508530.508549 1
- [Hae90] HAEERLI P.: Paint by numbers: Abstract image representations. *ACM SIGGRAPH Computer Graphics* 24, 3 (Aug. 1990), 207–214. doi: 10.1145/97879.97902 3
- [Hal03] HALPER N.: *Supportive Presentation for Computer Games*. PhD thesis, University of Magdeburg, Germany, 2003. urn:nbn:de:101:1-201010122864. 1
- [Hal14] HALLBERG J.: *Adolescents in a digital everyday environment*. PhD thesis, University of Gothenburg, Sweden, 2014. doi: 2077/36676 14
- [HE09] HEAD A., EISENBERG M.: Lessons learned: How college students seek information in the digital age. *SSRN* (December 2009). doi: 10.2139/ssrn.2281478 12
- [Her98] HERTZMANN A.: Painterly rendering with curved brush strokes of multiple sizes. In *Proc. SIGGRAPH* (1998), ACM, New York, pp. 453–460. doi: 10.1145/280814.280951 4, 5, 6
- [HL13] HALL P., LEHMANN A.-S.: Don’t measure—Appreciate! NPR seen through the prism of art history. In *Image and Video based Artistic Stylisation*, Rosin P., Collomosse J., (Eds.), vol. 42 of *Computational Imaging and Vision*. Springer, London, Heidelberg, 2013, ch. 16, pp. 333–351. doi: 10.1007/978-1-4471-4519-6_16 0

- [HMH*03] HALPER N., MELLIN M., HERRMANN C. S., LINNEWEBER V., STROTHOTTE T.: Psychology and non-photorealistic rendering: The beginning of a beautiful relationship. In *Proc. Mensch & Computer: Interaktion in Bewegung* (2003), Teubner Verlag, Stuttgart/Leipzig/Wiesbaden, pp. 277–286. doi: 10.1007/978-3-322-80058-9_28 1
- [HS14] HABERKAMP A., SCHMIDT T.: Enhanced visuomotor processing of phobic images in blood-injury-injection fear. *Journal of Anxiety Disorders* 28, 3 (Apr. 2014), 291–300. doi: 10.1016/j.janxdis.2014.02.001 2
- [HST13] HE K., SUN J., TANG X.: Guided image filtering. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 35, 6 (June 2013), 1397–1409. doi: 10.1109/TPAMI.2012.213 3
- [HWBS70] HARE R., WOOD K., BRITAIN S., SHADMAN J.: Autonomic responses to affective visual stimulation. *Psychophysiology* 7, 3 (Nov. 1970), 408–417. doi: 10.1111/j.1469-8986.1970.tb01766.x 2
- [IK12] INGLIS T. C., KAPLAN C. S.: Pixelating vector line art. In *Proc. Expressive* (2012), Eurographics Association, Aire-la-Ville, Switzerland, pp. 21–28. doi: 10.2312/PE/NPAR/NPAR12/021-028 3
- [Is13] ISENBERG T.: Evaluating and validating non-photorealistic and illustrative rendering. In *Image and Video based Artistic Stylization*, Rosin P., Collomosse J., (Eds.), vol. 42 of *Computational Imaging and Vision*. Springer, London, Heidelberg, 2013, ch. 15, pp. 311–331. doi: 10.1007/978-1-4471-4519-6_15 0
- [Is16] ISENBERG T.: Interactive NPAR: What type of tools should we create? In *Proc. NPAR* (2016), Eurographics Association, Goslar, Germany, pp. 89–96. doi: 10.2312/exp.20161067 0, 1
- [JAFF16] JOHNSON J., ALAHI A., FEI-FEI L.: Perceptual losses for real-time style transfer and super-resolution. In *Proc. ECCV* (2016), Springer International, Cham, Switzerland, pp. 694–711. doi: 10.1007/978-3-319-46475-6_43 13
- [Joh18] JOHNSTON L.: Whetting the appetite: What are the challenges for journalists covering Syria's conflict and the migration crisis using UGC? *Journal of Applied Journalism & Media Studies* 7, 1 (Mar. 2018), 177–195. doi: 10.1386/ajms.7.1.177_1 14
- [JS04] JIN R., SI L.: A study of methods for normalizing user ratings in collaborative filtering. In *Proc. SIGIR* (2004), ACM, New York, pp. 568–569. doi: 10.1145/1008992.1009124 6
- [Juk17] JUKES S.: *Affective journalism – Uncovering the affective dimension of practice in the coverage of traumatic news*. PhD thesis, Goldsmiths, University of London, UK, Apr. 2017. doi: 10.25602/GOLD.00020635 14
- [JWK12] JUDD C. M., WESTFALL J., KENNY D. A.: Treating stimuli as a random factor in social psychology: A new and comprehensive solution to a pervasive but largely ignored problem. *Journal of Personality and Social Psychology* 103, 1 (July 2012), 54–69. doi: 10.1037/a0028347 14
- [KA13] KRZYWINSKI M., ALTMAN N.: Points of significance: Error bars. *Nature Methods* 10, 10 (Oct. 2013), 921–922. doi: 10.1038/nmeth.2659 9
- [KCWI13] KYPRIANIDIS J. E., COLLOMOSSE J., WANG T., ISENBERG T.: State of the “art”: A taxonomy of artistic stylization techniques for images and video. *IEEE Transactions on Visualization and Computer Graphics* 19, 5 (May 2013), 866–885. doi: 10.1109/TVCG.2012.160 0, 3
- [KD08] KYPRIANIDIS J. E., DÖLLNER J.: Image abstraction by structure adaptive filtering. In *Proc. EG UK—Theory and Practice of Computer Graphics* (2008), Eurographics Association, Goslar, Germany, pp. 51–58. doi: 10.2312/LocalChapterEvents/TPCG/TPCG08/051-058 0, 3, 5, 6, 10, 13
- [KG13] KIRBY K. N., GERLANC D.: BootES: An R package for bootstrap confidence intervals on effect sizes. *Behavior Research Methods* 45, 4 (Dec. 2013), 905–927. doi: 10.3758/s13428-013-0330-5 9
- [KH18] KLAUITTER E., HARGITTAI E.: “I went home to Google”: How users assess the credibility of online health information. In *eHealth: Current Evidence, Promises, Perils and Future Directions*, Hale T. M., Chou W.-Y. S., Cotten S. R., Khilnani A., (Eds.), vol. 15. Emerald Publishing Limited, 2018, ch. 1, pp. 11–41. doi: 10.1108/S2050-20602018000015001 12
- [KHEK76] KUWAHARA M., HACHIMURA K., EIHO S., KINOSHITA M.: Processing of RI-angiocardiac images. In *Digital Processing of Biomedical Images* (1976), Plenum Press, New York, pp. 187–202. doi: 10.1007/978-1-4684-0769-3_13 4
- [Kil17] KILL BILL WIKI CONTRIBUTORS: Crazy 88, October 2017. URL: http://killbill.wikia.com/wiki/Crazy_88. 3
- [KJDL09] KIM Y., JANG C., DEMOUTH J., LEE S.: Robust color-to-gray via nonlinear global mapping. *ACM Transactions on Graphics* 28, 5 (Dec. 2009), 161:1–161:4. doi: 10.1145/1618452.1618507 3
- [KK11] KYPRIANIDIS J. E., KANG H.: Image and video abstraction by coherence-enhancing filtering. *Computer Graphics Forum* 30, 2 (Apr. 2011), 593–602. doi: 10.1111/j.1467-8659.2011.01882.x 4, 5, 6, 10, 13
- [KKD09] KYPRIANIDIS J. E., KANG H., DÖLLNER J.: Image and video abstraction by anisotropic Kuwahara filtering. *Computer Graphics Forum* 28, 7 (Oct. 2009), 1955–1963. doi: 10.1111/j.1467-8659.2009.01574.x 4
- [KL08] KANG H., LEE S.: Shape-simplifying image abstraction. *Computer Graphics Forum* 27, 7 (Oct. 2008), 1773–1780. doi: 10.1111/j.1467-8659.2008.01322.x 4, 5, 6
- [KLC07] KANG H., LEE S., CHUI C. K.: Coherent line drawing. In *Proc. NPAR* (2007), ACM, New York, pp. 43–50. doi: 10.1145/1274871.1274878 5
- [KLC09] KANG H., LEE S., CHUI C. K.: Flow-based image abstraction. *IEEE Transactions on Visualization and Computer Graphics* 15, 1 (Jan./Feb. 2009), 62–76. doi: 10.1109/TVCG.2008.81 3
- [KS10] KASS M., SOLOMON J.: Smoothed local histogram filters. *ACM Transactions on Graphics* 29, 4 (July 2010), 100:1–100:10. doi: 10.1145/1833351.1778837 5
- [KWW77] KLORMAN R., WEISSBERG R. P., WIESENFIELD A. R.: Individual differences in fear and autonomic reactions to affective stimulation. *Psychophysiology* 14, 1 (Jan. 1977), 45–51. doi: 10.1111/j.1469-8986.1977.tb01154.x 2
- [Kyp11] KYPRIANIDIS J. E.: Image and video abstraction by multi-scale anisotropic Kuwahara filtering. In *Proc. NPAR* (2011), ACM, New York, pp. 55–64. doi: 10.1145/2024676.2024686 4, 5, 6, 10
- [LBC97] LANG P. J., BRADLEY M. M., CUTHBERT B. N.: International affective picture system (IAPS): Technical manual and affective ratings. *NIMH Center for the Study of Emotion and Attention* (1997), 39–58. 7, 8
- [LGBH93] LANG P. J., GREENWALD M. K., BRADLEY M. M., HAMM A. O.: Looking at pictures: Affective, facial, visceral, and behavioral reactions. *Psychophysiology* 30, 3 (May 1993), 261–273. doi: 10.1111/j.1469-8986.1993.tb03352.x 2
- [LKPS14] LIVINGSTONE S., KIRWIL L., PONTE C., STAKSRUD E.: In their own words: What bothers children online? *European Journal of Communication* 29, 3 (Mar. 2014), 271–288. doi: 10.1177/0267323114521045 12, 14
- [LM92] LUMLEY M. A., MELAMED B. G.: Blood phobics and nonphobics: Psychological differences and affect during exposure. *Behaviour Research and Therapy* 30, 5 (Sept. 1992), 425–434. doi: 10.1016/0005-7967(92)90026-D 2
- [LSG*18] LEWIS L., SOMERS J. M., GUY R., WATCHIRS-SMITH L., SKINNER S. R.: “I see it everywhere”: Young Australians unintended exposure to sexual content online. *Sexual health* 15 (2018), 335–341. doi: 10.1071/SH17132 12, 14
- [LVPI18] LAWONN K., VIOLA I., PREIM B., ISENBERG T.: A survey of surface-based illustrative rendering for visualization. *Computer Graphics Forum* 37, 6 (Sept. 2018), 205–234. doi: 10.1111/cgf.13322 1
- [Mad10] MADDISON S.: Online obscenity and myths of freedom: Dangerous images, child porn and neoliberalism. In *Porn.com: Making sense of online pornography*, Attwood F., (Ed.). Peter Lang, New York, 2010. 14
- [MALI10] MARTÍN D., ARROYO G., LUZÓN M. V., ISENBERG T.: Example-based stippling using a scale-dependent grayscale process. In *Proc. NPAR* (2010), ACM, New York, pp. 51–61. doi: 10.1145/1809939.1809946 4, 5

- [MALI11] MARTÍN D., ARROYO G., LUZÓN M. V., ISENBERG T.: Scale-dependent and example-based stippling. *Computers & Graphics* 35, 1 (Feb. 2011), 160–174. doi: 10.1016/j.cag.2010.11.006 4, 5, 6
- [MB10] MOFADDEL M., BAKHEET S.: Adult image content filtering: A statistical method based on multi-color skin modeling. In *Proc. ISSPIT* (2010), IEEE, Piscataway, NJ, USA, pp. 366–370. doi: 10.1109/ISSPIT.2010.5711812 14
- [MBB12] McDONNELL R., BREIDT M., BÜLTHOFF H. H.: Render me real? Investigating the effect of render style on the perception of animated virtual humans. *ACM Transactions on Graphics* 31, 4 (July 2012), 91:1–91:11. doi: 10.1145/2185520.2185587 1
- [MH80] MARR D., HILDRETH E.: Theory of edge detection. *Proceedings of the Royal Society B* 207, 1167 (Feb. 1980), 187–217. doi: 10.1098/rspb.1980.0020 5
- [MMA*16] MARTELLOZZO E., MONAGHAN A., ADLER J. R., DAVIDSON J., LEYVA R., HORVATH M. A.: "...I wasn't sure it was normal to watch it..." A quantitative and qualitative examination of the impact of online pornography on the values, attitudes, beliefs and behaviours of children and young people. Commissioned research report, Middlesex University, London, June 2016. doi: 10.6084/m9.figshare.3382393 14
- [MML11] MANDRYK R. L., MOULD D., LI H.: Evaluation of emotional response to non-photorealistic images. In *Proc. NPAR* (2011), ACM, New York, pp. 7–16. doi: 10.1145/2024676.2024678 1, 5
- [MML12] MOULD D., MANDRYK R. L., LI H.: Emotional response and visual attention to non-photorealistic images. *Computers & Graphics* 36, 6 (Oct. 2012), 658–672. doi: 10.1016/j.cag.2012.03.039 1
- [Mou12] MOULD D.: Texture-preserving abstraction. In *Proc. NPAR* (2012), Eurographics Association, Goslar, Germany, pp. 75–82. doi: 10.2312/PE/NPAR/NPAR12/075-082 5
- [Mou14] MOULD D.: Authorial subjective evaluation of non-photorealistic images. In *Proc. NPAR* (2014), ACM, New York, pp. 49–56. doi: 10.1145/2630397.2630400 0
- [MR16] MOULD D., ROSIN P. L.: A benchmark image set for evaluating stylization. In *Proc. NPAR* (2016), The Eurographics Association, Goslar, Germany, pp. 11–20. doi: 10.2312/exp.20161059 14
- [MR17] MOULD D., ROSIN P. L.: Developing and applying a benchmark for evaluating image stylization. *Computers & Graphics* 67 (Oct. 2017), 58–76. doi: 10.1016/j.cag.2017.05.025 14
- [MS16] MAHMOODI M. R., SAYEDI S. M.: A comprehensive survey on human skin detection. *International Journal of Image, Graphics & Signal Processing* 8, 5 (May 2016), 1–35. doi: 10.5815/ijgsp.2016.05.01 13
- [MŽJG14] MARCHEWKA A., ŻURAWSKI Ł., JEDNORÓG K., GRABOWSKA A.: The Nencki Affective Picture System (NAPS): Introduction to a novel, standardized, wide-range, high-quality, realistic picture database. *Behavior Research Methods* 46, 2 (June 2014), 596–610. doi: 10.3758/s13428-013-0379-1 7, 8
- [MZZW15] MA K., ZHAO T., ZENG K., WANG Z.: Objective quality assessment for color-to-gray image conversion. *IEEE Transactions on Image Processing* 24, 12 (Dec. 2015), 4673–4685. doi: 10.1109/TIP.2015.2460015 3
- [NM08] NICHOLS A. L., MANER J. K.: The good-subject effect: Investigating participant demand characteristics. *The Journal of General Psychology* 135, 2 (2008), 151–166. doi: 10.3200/GENP.135.2.151-166 14
- [NvN07] NEUMANN L., ČADÍK M., NEMCSICS A.: An efficient perception-based adaptive color to gray transformation. In *Proc. CAE* (2007), Eurographics Association, Goslar, Germany, pp. 73–80. doi: 10.2312/COMPAESTH/COMPAESTH07/073-080 3
- [OBMR12] OWENS E. W., BEHUN R. J., MANNING J. C., REID R. C.: The impact of internet pornography on adolescents: A review of the research. *Sexual Addiction & Compulsivity* 19, 1–2 (Apr. 2012), 99–122. doi: 10.1080/10720162.2012.660431 14
- [OCMP10] OLATUNJI B. O., CISLER J., MCKAY D., PHILLIPS M. L.: Is disgust associated with psychopathology? Emerging research in the anxiety disorders. *Psychiatry Research* 175, 1–2 (Jan. 2010), 1–10. doi: 10.1016/j.psychres.2009.04.007 2
- [OHMD08] OLATUNJI B. O., HAIDT J., MCKAY D., DAVID B.: Core, animal reminder, and contamination disgust: Three kinds of disgust with distinct personality, behavioral, physiological, and clinical correlates. *Journal of Research in Personality* 42, 5 (Oct. 2008), 1243–1259. doi: 10.1016/j.jrp.2008.03.009 2
- [ÖSL84] ÖST L.-G., STERNER U., LINDAHL I.-L.: Physiological responses in blood phobics. *Behaviour Research and Therapy* 22, 2 (1984), 109–117. doi: 10.1016/0005-7967(84)90099-8 2
- [PAM*17] PEREZ M., AVILA S., MOREIRA D., MORAES D., TESTONI V., VALLE E., GOLDENSTEIN S., ROCHA A.: Video pornography detection through deep learning techniques and motion information. *Neurocomputing* 230 (Mar. 2017), 279–293. doi: 10.1016/j.neucom.2016.12.017 14
- [PB14] PREIM B., BOTHA C.: Illustrative medical visualization. In *Visual Computing for Medicine*, 2nd ed. Morgan Kaufmann, Boston, 2014, ch. 12, pp. 451–508. doi: 10.1016/B978-0-12-415873-3.00012-2 1
- [PDF14] PERIN C., DRAGICEVIC P., FEKETE J.-D.: Revisiting Bertin matrices: New interactions for crafting tabular visualizations. *IEEE Transactions on Visualization and Computer Graphics* 20, 12 (Dec. 2014), 2082–2091. doi: 10.1109/TVCG.2014.2346279 6
- [PE10] POLASTRO M. D. C., ELEUTERIO P. M. D. S.: NuDetective: A forensic tool to help combat child pornography through automatic nudity detection. In *Proc. Workshops on Database and Expert Systems Applications* (2010), IEEE, Piscataway, NJ, USA, pp. 349–353. doi: 10.1109/DEXA.2010.74 14
- [PET08] PETA: Why does PETA use graphic imagery? *PETA* (2008). URL: <https://www.peta.org/blog/peta-use-graphic-imagery/>. 12
- [PHWF01] PRAUN E., HOPPE H., WEBB M., FINKELSTEIN A.: Real-time hatching. In *Proc. SIGGRAPH* (2001), ACM, New York, pp. 581–586. doi: 10.1145/383259.383328 4, 5, 6
- [Pos14] POSETTI J.: Is social media to blame for the increase of graphic images in media? *MediaShift* (2014). URL: <http://mediashift.org/2014/08/is-social-media-to-blame-for-the-increase-in-graphic-images-in-media/>. 12
- [PPC07] PAPARI G., PETKOV N., CAMPISI P.: Artistic edge and corner enhancing smoothing. *IEEE Transactions on Image Processing* 16, 10 (Oct. 2007), 2449–2662. doi: 10.1109/TIP.2007.903912 4
- [Pra07] PRATT W. K.: *Digital Image Processing: PIKS Scientific Inside*, 4th ed. John Wiley & Sons, Inc., Hoboken, NJ, USA, 2007. doi: 10.1002/0470097434 3, 5
- [PSL14] PLATZER C., STUETZ M., LINDORFER M.: Skin sheriff: A machine learning solution for detecting explicit images. In *Proc. SFCS* (2014), ACM, New York, pp. 45–56. doi: 10.1145/2598918.2598920 13
- [PV09] PETER J., VALKENBURG P. M.: Adolescents' exposure to sexually explicit internet material and notions of women as sex objects: Assessing causality and underlying processes. *Journal of Communication* 59, 3 (Sept. 2009), 407–433. doi: 10.1111/j.1460-2466.2009.01422.x 14
- [PV16] PETER J., VALKENBURG P. M.: Adolescents and pornography: A review of 20 years of research. *The Journal of Sex Research* 53, 4–5 (2016), 509–531. PMID: 27105446. doi: 10.1080/00224499.2016.1143441 14
- [RBGV08] RAUTEK P., BRUCKNER S., GRÖLLER E., VIOLA I.: Illustrative visualization: New technology or useless tautology? *ACM SIGGRAPH Computer Graphics* 42, 3 (Aug. 2008), 4:1–4:8. doi: 10.1145/1408626.1408633 1
- [RC13] ROSIN P., COLLOMOSSE J. (Eds.): *Image and Video based Artistic Stylisation*, vol. 42 of *Computational Imaging and Vision*. Springer, London, Heidelberg, 2013. doi: 10.1007/978-1-4471-4519-6 0
- [RH08] ROHRMANN S., HOPP H.: Cardiovascular indicators of disgust. *International Journal of Psychophysiology* 68, 3 (June 2008), 201–208. doi: 10.1016/j.jpsycho.2008.01.011 2

- [RMB*17] ROSIN P. L., MOULD D., BERGER I., COLLOMOSSE J., LAI Y.-K., LI C., LI H., SHAMIR A., WAND M., WANG T., WINNEMÖLLER H.: Benchmarking non-photorealistic rendering of portraits. In *Proc. NPAR* (2017), ACM, New York, pp. 11:1–11:12. doi: 10.1145/3092919.3092921 14
- [RMSS15] RENNIS L., MCNAMARA G., SEIDEL E., SHNEYDERMAN Y.: Google it!: Urban community college students' use of the internet to obtain selfcare and personal health information. *College Student Journal* 49, 3 (Fall 2015), 414–426. 12
- [RRH*02] RICHARDSON C. R., RESNICK P. J., HANSEN D. L., DERRY H. A., RIDEOUT V. J.: Does pornography-blocking software block access to health information on the internet? *JAMA* 288, 22 (Dec. 2002), 2887–2894. doi: 10.1001/jama.288.22.2887 12
- [Sal02] SALESIN D. H.: Non-photorealistic animation & rendering: 7 grand challenges. Keynote talk at NPAR, June 2002. 0
- [SBC06] SHUGRINA M., BETKE M., COLLOMOSSE J.: Empathic painting: Interactive stylization through observed emotional state. In *Proc. NPAR* (2006), ACM, New York, pp. 87–96. doi: 10.1145/1124728.1124744 2
- [SBPG03] SKINNER H., BISCOPE S., POLAND B., GOLDBERG E.: How adolescents use technology for health information: Implications for health professionals from focus group studies. *Journal of Medical Internet Research* 5, 4 (Dec 2003), e32:1–e32:10. doi: 10.2196/jmir.5.4.e32 12
- [SDN18] SMITH R. J., DREVO S., NEWMAN E.: Covering traumatic news stories: Factors associated with post-traumatic stress disorder among journalists. *Stress and Health* 34, 2 (Apr. 2018), 218–226. doi: 10.1002/smi.2775 14
- [She18] SHEN C.-X.: Does school-related internet information seeking improve academic self-efficacy? The moderating role of internet information seeking styles. *Computers in Human Behavior* 86 (Sept. 2018), 91–98. doi: 10.1016/j.chb.2018.04.035 12
- [SID17] SEMMO A., ISENBERG T., DÖLLNER J.: Neural style transfer: A paradigm shift for image-based artistic rendering? In *Proc. NPAR* (2017), ACM, New York, pp. 5:1–5:13. doi: 10.1145/3092919.3092920 13
- [SIMC07] SCHWARZ M., ISENBERG T., MASON K., CARPENDALE S.: Modeling with rendering primitives: An interactive non-photorealistic canvas. In *Proc. NPAR* (2007), ACM, New York, pp. 15–22. doi: 10.1145/1274871.1274874 3
- [SL11] SAURO J., LEWIS J. R.: When designing usability questionnaires, does it hurt to be positive? In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (2011), CHI '11, ACM, New York, NY, USA, pp. 2215–2224. doi: 10.1145/1978942.1979266 8
- [SLKD16] SEMMO A., LIMBERGER D., KYPRIANIDIS J. E., DÖLLNER J.: Image stylization by interactive oil paint filtering. *Computers & Graphics* 55 (Apr. 2016), 157–171. doi: 10.1016/j.cag.2015.12.001 4, 5, 6
- [SISW*02] SCHIENLE A., STARK R., WALTER B., BLECKER C., OTT U., KIRSCH P., SAMMER G., VAITL D.: The insula is not specifically involved in disgust processing: An fMRI study. *NeuroReport* 13, 16 (Nov. 2002), 2023–2026. doi: 10.1097/00001756-200211150-00006 2
- [SLTM08] SMITH K., LANDES P.-E., THOLLOT J., MYSZKOWSKI K.: Apparent greyscale: A simple and fast conversion to perceptually accurate images and video. *Computer Graphics Forum* 27, 2 (Apr. 2008), 193–200. doi: 10.1111/j.1467-8659.2008.01116.x 3, 4, 5
- [SLW*02] SAWCHUK C. N., LOHR J. M., WESTENDORF D. H., MEUNIER S. A., TOLIN D. F.: Emotional responding to fearful and disgusting stimuli in specific phobias. *Behaviour Research and Therapy* 40, 9 (Sept. 2002), 1031–1046. doi: 10.1016/S0005-7967(01)00093-6 1, 2
- [SP14] SHIM J. W., PAUL B. M.: The role of anonymity in the effects of inadvertent exposure to online pornography among young adult males. *Social Behavior and Personality: An International Journal* 42, 5 (June 2014), 823–834. doi: 10.2224/sbp.2014.42.5.823 14
- [SS02] STROTHOTTE T., SCHLECHTWEIG S.: *Non-Photorealistic Computer Graphics. Modeling, Animation, and Rendering*. Morgan Kaufmann Publishers, San Francisco, 2002. doi: 10.1016/B978-1-55860-787-3.50019-0 1
- [SSRL96] SCHUMANN J., STROTHOTTE T., RAAB A., LASER S.: Assessing the effect of non-photorealistic rendered images in CAD. In *Proc. CHI* (1996), ACM, New York, pp. 35–42. doi: 10.1145/238386.238398 1
- [SWR08] SKOPELJA E. N., WHIPPLE E. C., RICHWINE P.: Reaching and teaching teens: Adolescent health literacy and the internet. *Journal of Consumer Health on the Internet* 12, 2 (Oct. 2008), 105–118. doi: 10.1080/15398280802121406 12
- [TBO*17] THIRD A., BELLEROSSE D., OLIVEIRA J., LALA G., THEAKSTONE G.: Young and online: Children's perspectives on life in the digital age. *The State of the World's Children 2017 Companion Report* (2017). 14
- [TSL97] TOLIN D. F., LOHR J. M., SAWCHUK C. N., LEE T. C.: Disgust and disgust sensitivity in blood-injection-injury and spider phobia. *Behaviour Research and Therapy* 35, 10 (Oct. 1997), 949–953. doi: 10.1016/S0005-7967(97)00048-X 1, 2
- [TM98] TOMASI C., MANDUCHI R.: Bilateral filtering for gray and color images. In *Proc. ICCV* (1998), IEEE Computer Society, Los Alamitos, pp. 839–846. doi: 10.1109/CCV.1998.710815 3, 4, 5, 6, 10
- [Too14] TOOTH R.: Graphic content: When photographs of carnage are too upsetting to publish. *The Guardian* (July 2014). URL: <https://www.theguardian.com/world/2014/jul/23/graphic-content-photographs-too-upsetting-to-publish-gaza-mh17-ukraine>. 2, 12
- [TVT18a] TVTROPES CONTRIBUTORS: Adjustable censorship, March 2018. URL: <http://tvtropes.org/pmwiki/pmwiki.php/Main/AdjustableCensorship>. 3
- [TVT18b] TVTROPES CONTRIBUTORS: Black blood, March 2018. URL: <http://tvtropes.org/pmwiki/pmwiki.php/Main/BlackBlood>. 3
- [Twi18] TWITTER: Twitter media policy, October 2018. URL: <https://help.twitter.com/en/rules-and-policies/media-policy>, version of October 2018. 12, 13
- [VCSE*06] VIOLA I., COSTA SOUSA M., EBERT D. S., ANDREWS B., GOOCH B., TIETJEN C.: Illustrative visualization for medicine and science. In *Eurographics Tutorials* (2006), Eurographics Association, Goslar, Germany, pp. 1061–1200. doi: 10.2312/egt.20061068 1
- [VI18] VIOLA I., ISENBERG T.: Pondering the concept of abstraction in (illustrative) visualization. *IEEE Transactions on Visualization and Computer Graphics* 24, 9 (Sept. 2018), 2573–2588. doi: 10.1109/TVCG.2017.2747545 1
- [vSSC13] VAN SONDEREN E., SANDERMAN R., COYNE J. C.: Ineffectiveness of reverse wording of questionnaire items: let's learn from cows in the rain. *PloS one* 8, 7 (07 2013), e68967. doi: 10.1371/journal.pone.0068967 8
- [Wei99] WEICKERT J.: Coherence-enhancing diffusion filtering. *International Journal of Computer Vision* 31, 2–3 (Apr. 1999), 111–127. doi: 10.1023/A:1008009714131 3, 10, 13
- [Wik03] WIKIPEDIA CONTRIBUTORS: Wikipedia:content disclaimer, 2003. Revision 02/14/2016, URL: https://en.wikipedia.org/wiki/Wikipedia:Content_disclaimer. 12
- [Wik10a] WIKIPEDIA CONTRIBUTORS: QA Wikimedia Commons images review, 2010. URL: https://foundation.wikimedia.org/wiki/QA_Wikimedia_Commons_images_review_May_2010. 2, 12, 13
- [Wik10b] WIKIPEDIA CONTRIBUTORS: Wikipedia:sexual content/FAQ, 2010. URL: https://en.wikipedia.org/wiki/Wikipedia:Sexual_content/FAQ. 2, 12
- [WKG*03] WRASE J., KLEIN S., GRUESSER S. M., HERMANN D., FLOR H., MANN K., BRAUS D. F., HEINZ A.: Gender differences in the processing of standardized emotional visual stimuli in humans: A functional magnetic resonance imaging study. *Neuroscience Letters* 348, 1 (Sept. 2003), 41–45. doi: 10.1016/S0304-3940(03)00565-2 2
- [WKO12] WINNEMÖLLER H., KYPRIANIDIS J. E., OLSEN S. C.: XDoG: An extended difference-of-Gaussians compendium including advanced image stylization. *Computers & Graphics* 36, 6 (Oct. 2012), 740–753. doi: 10.1016/j.cag.2012.03.004 4, 5, 6, 10

- [WOG06] WINNEMÖLLER H., OLSEN S. C., GOOCH B.: Real-time video abstraction. *ACM Transactions on Graphics* 25, 3 (July 2006), 1221–1226. doi: 10.1145/1141911.1142018 3
- [WTK15] WRIGHT P. J., TOKUNAGA R. S., KRAUS A.: A meta-analysis of pornography consumption and actual acts of sexual aggression in general population studies. *Journal of Communication* 66, 1 (Feb. 2015), 183–205. doi: 10.1111/jcom.12201 14
- [WTS15] WALKER S., TEMPLE-SMITH M., HIGGS P., SANCI L.: ‘It’s always just there in your face’: Young people’s views on porn. *Sexual Health* 12, 3 (June 2015), 200–206. doi: 10.1071/SH14225 14
- [WWF*14] WANG M., WANG B., FEI Y., QIAN K., WANG W., CHEN J., YONG J.-H.: Towards photo watercolorization with artistic verisimilitude. *IEEE Transactions on Visualization and Computer Graphics* 20, 10 (Feb. 2014), 1451–1460. doi: 10.1109/TVCG.2014.2303984 4, 5, 6
- [XLXJ11] XU L., LU C., XU Y., JIA J.: Image smoothing via L_0 gradient minimization. *ACM Transactions on Graphics* 30, 6 (Dec. 2011), 174:1–174:12. doi: 10.1145/2070781.2024208 5
- [Yan10] YAN Y. Y.: Online health information seeking behavior in Hong Kong: An exploratory study. *Journal of Medical Systems* 34, 2 (Apr 2010), 147–153. doi: 10.1007/s10916-008-9226-9 12
- [Zel18] ZELDAWIKI CONTRIBUTORS: Controversy in The Legend of Zelda series, March 2018. URL: https://zelda.gamepedia.com/Controversy_in_The_Legend_of_Zelda_Series. 3
- [ZFB17] ZUO H., FAN H., BLASCH E., LING H.: Combining convolutional and recurrent neural networks for human skin detection. *IEEE Signal Processing Letters* 24, 3 (March 2017), 289–293. doi: 10.1109/LSP.2017.2654803 13